Universidad Autónoma de Madrid
Departamento de Qumica

Université Evry Val d'Essonne
LAMBE

Fragmentation dynamics of biomolecules in gas phase and water environment

A thesis submitted for the degree of
Doctor

Candidate :
Pablo López Tarifa

Supervisors :
Manuel Alcamí Pertejo
Marie-Pierre Gaigeot

Madrid, November 2010

Resumen

Hoy en día, el empleo de iones muy energéticos y altamente cargados se ha convertido en una exitosa alternativa a los tradicionales rayos-X en el tratamiento del cáncer. Cuando son incididos sobre tejidos vivos, dichos iones presentan una región en la cual su depósito de energía es máximo (región de Bragg); permitiendo así una mayor localización de la dósis a una determinda profundidad de penetración. Desafortunadamente la formación de las especies qumicas resultantes de dicha interacción (y que pueden llegar a inducir la muerte celular) no pueden ser establecidos experimentalmente ya que ocurren a una escala de tiempo muy reducida ($\approx$ 1 - 100 fs).

Las simulaciones basadas en Dinámica Molecular (MD) han demostrado ser la herramienta idónea para el estudio de tales mecanismos, sin embargo, hasta ahora muy pocos trabajos se han centrado en estos procesos primarios. En la presente tesis hemos aplicado una combinación de MD basada en la Teoría del Funcional de la Densidad Tiempo-Dependiente (TD-DFT) y en la approximación de Born-Oppenheimer (BO), con el fin de investigar la fragmentación de especies doblemente cargadas tales como uracilo, pirimidina y pequeños agregados de agua, generados mediante la ionizaciones de capas internas.

El estado inicial doblemente cargado es preparado mediante la sustracción de dos electrones de un orbital específico de Kohn-Sham o Wannier de la especie neutra. La densidad electrónica evoluciona según la llamada aproximación Ehrenfest MD donde una superficie electrónica promedio, calculada al nivel TD-DFT (mediante la propagación tiempo-depenediente de las ecuaciones de Kohn-Sham) conduce el movimiento nuclear. Las fuerza que actúan sobre cada átomo son calculadas mediante el teorema de Hellmann-Feynman. Cuando la dinámica es esencialmente adiabática, la superfiecie de energía potencial es la correspodiente al estado fundamental y en cosecuencia, la dinámica es extendida mediante una BO MD hasta un tiempo total de $\approx$ 100 fs.

Los resultados obtenidos para la fragmentación de uracilo^{2+} en fase gas están en buen acuerdo con las coincidencias ion-ion medidas en los experimentos de colisión de uracilo con protones de 100 keV. Orbitales de misma engía y/o localizados en los mismos enlaces pueden llevar a muy diferenetes fragmentaciones, mostrando así la importancia del medio químico intramolecular. Los fragmentos observados casi nunca corresponden con los caminos de disociación más favorables, lo que pone de manifiesto la importancia de los efectos dinámicos que ocurren en los primeros fs después de la ionización. Cuando la biomolécula se encuentra hidratada, la ionización de los mismos orbitales de Kohn-Sham que los propios de la molécula aislada lleva a diferentes patrones de fragmentación. El disolvente representa un papel activo en el proceso ya que no sólo evita la movilidad de la carga en el anillo durante los primeros femtosegundos, sino que también actúa tomando carga positiva de la biomolécula y llevándola a capas de hidratación más externas. Para evaluar los efectos indirectos de la radiación, hemos supuesto también la doble ionización de una molécula de agua vecina del uracilo. Así, ionizaciones tanto de una molécula de agua como del propio uracilo llevan a la formación de oxígeno atómico como consecuencia directa de la explosión Coulombiana.

El mismo método ha sido empleado en el estudio de la disociación de agregados de agua por

impacto de iones altamente cargados. Nuestros resultados respaldan la alta movilidad de carga sugerida por los espectros de masa en coicidencia. Por otra parte y de forma preliminar, el estudio TD-DFT MD también ha sido extendido al ion pirmimidina^{2+} generado mediante experimentos de coincidencia foto-electrón-ion en fase gas. La combinación de ambos, experimento y teória, apunta a una asignación inequívoca entre el orbital ionizado y la señal en el espectro de masas producida.

Abstract

Nowadays, swift heavy ion beams have become a successfully alternative to traditional X-Ray radiation for cancer treatment. Their high energy loss region (Bragg peak) allows to target the maximum deposited energy at a selected body depth. Unfortunately, formation of the chemical species resulting from the irradiation, which might lead to the cell death, cannot be established by experiments since these processes occur in a very short time scale ($\approx$ 1 - 100 fs).

Molecular Dynamics (MD) simulations have demonstrated to be a well suited tool to unravel such mechanisms, however, so far only a few studies have focused on these primary processes. In the present thesis we apply a combination of Time-Dependent Density Functional Theory (TD-DFT) MD and Born-Oppenheimer (BO) MD methods to investigate the fragmentation of doubly-charged uracil, pyrimidine and small water clusters arisen from inner shell ionizations.

The initial doubly charged state is prepared by removing two electrons from a specific Kohn-Sham or Wannier orbital of the neutral species. The electronic density evolves within the so-called Ehrenfest MD approximation where the mean field potential energy surface driving the nuclear dynamics is computed at the TD-DFT level (by propagation of the time-dependent Kohn-Sham equations), and nuclear forces are computed using the Hellmann-Feynman theorem. When the dynamics becomes essentially adiabatic, the system is switched to ground state BO MD and is extended until total times next to $\approx$ 100fs.

Results for uracil^{2+} fragmention in gas phase are in good agreement with ion/ion coincidence measurements performed in uracil collisions with protons of 100 keV. Orbitals of similar energy and/or localized in similar bonds can lead to very different fragmentation patterns, thus showing the importance of the intramolecular chemical environment. The observed fragments almost never correspond to the energetically most favorable dissociation paths, which is due to dynamical effects occurring in the first few fs after electron removal. When the biomolecule is immersed in water, ionization from equivalent Kohn-Sham orbitals than in gas phase lead to different fragmentation patterns. The solvent is revealed as an active particpant in the dissociation process since it not only prevents the charge mobility of the biomolecule during the first femtoseconds of dynamics, but also drains positive charge from the uracil to outer hydratation shells. To evaluate the indirect effect of the radiation, double ionizations of a sourrounding water molecule of uracil are also considered. Ionizations either from one water molecule or the uracil lead to the formation of an atomic oxygen as a direct consequence of the molecule Coulomb explosion.

Same techniques have been also applied to the study of dissociative ionization of water clusters by heavy impact ion. Our results support the high charge mobility suggested by the mass spectrum coincidences. Preliminary, TD-DFT molecular dynamics are also extended to investigate fragmentation of doubly-charged gas-phase pyrimidine produced by photoelectron-ion coincidence experiments. Combination of both, experimental and theoretical time-dependent tools, is aimed to unambiguously identify the correspondence between the orbital ionized and the mass spectra signals obtained.

Contents

Part I

Introduction

Chapter 1

Introduction

1.1 Introduction

Though irradiation of malignant tumours by high energy protons or heavy ions has become a promising treatment in cancer therapy in the past few years, the physical and chemical processes involved in the induction of DNA damage by ionizing radiation are not completely elucidated. Biological damages are nowadays considered the results of attacks of DNA molecules that are classified as either direct or indirect effects [1]. Direct effects are usually considered as arising from the energy released inside DNA molecules within ionization tracks. Induced biological damages, on the other hand, are to a large extent recognized as to be due to the action of Reactive Oxygen Species (ROS) that originate, for instance, from the radiolysis of water molecules in the vicinity of DNA. The further chemical reactivity that can occur with DNA in this latter case is the so-called indirect effect.

The frontier between direct and indirect effects is however a puzzling question, since hydration shells are part of DNA molecules and as such also govern their conformations [2]. Numerous studies have suggested that ionization of these tightly bound water molecules results in damage to the DNA due to charge transfer, a phenomenon called the quasi-direct effect [3, 4]. The fact that no hydroxyl radicals and trapped electrons were detected below about 10 and 21 water molecules per nucleotide, respectively, has been presented as evidence to this effect [5, 6].

However, it is convenient to localize all of these events not only in space but also in time. As it is shown in Fig.(1.1), phenomena following irradiation are classified according into stages: The *physical stage* that lasts only about 10^{-15} s, and during which the ion track particles (and all the secondary electrons that they generate) are slowed-down through the excitation and the ionization of all molecules the ion-track finds on its way. Afterwards, there is a sequence of physical-chemical events that can lead DNA damage and commonly gathered in the so-called *physical-chemical stage*. During this period, that lasts from a few fs (femto-second) to the scale of seconds, there is time to the fragmentation of the ionized molecules due to the Coulobmian explosions, formation of the first radicals and the diffusion and recombination of these species to generate new ones. If the cell is damaged there is a very complex enzymatic process trying to fix it. If it is not the case, the cell can induce its own death (apoptosis) or if the damage remains, it can lead to the DNA mutation and cancer in a major time scale (days or years).

Our work focuses on the theoretical characterization of the primary events (10^{15} to 10^{12} s) subsequent to the ionization of molecules within ionization tracks of swift heavy ions. Our objective is to characterize the chemical products that result from the Coulomb explosion of the ionized molecules, and more particularly decipher the products arising from direct ionization of DNA from the products arising from ionization of the surrounding water hydration shells.

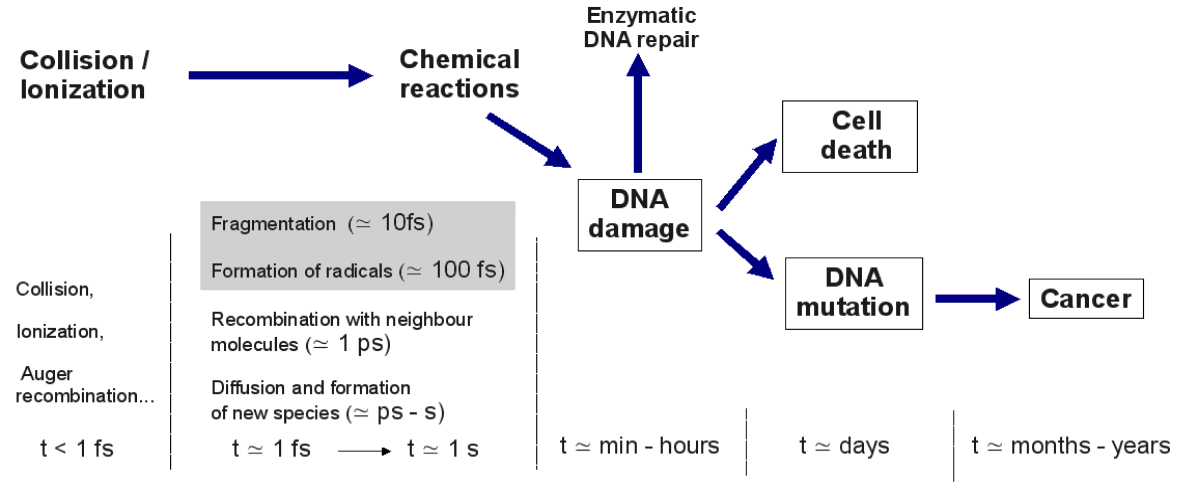


Figure 1.1: Schematic time-diagram of molecular damage. Highlighted in grey the time scale studied.

Investigation of the early processes that take place after ionization of cells elementary constituents (such as water molecules, nucleotides and proteins) remains a challenge from both experimental and theoretical points of view, but is required for a full understanding of the mechanisms of radiation damage in cells. The advent of femtosecond time-resolved laser spectroscopy has enabled the investigation of early charge-transfer processes and elementary radical reactions in molecular liquids and solutions [7]. These femtolysis experimental results clearly demonstrate the inadequacy of classical analysis (framework of stochastic treatment with master equations of the diffusion) of short range coupling between non-equilibrium electron and nascent water decomposition products. See Ref. [8] and references therein. Moreover, Gauduel et al. have observed ultrafast one-electron reduction of oxidized pyridine nucleotides [9] and cystamine [10] by prehydrated electrons. These findings and more recent real-time observation of dissociative electron transfer reactions of prehydrated electrons with DNA nucleotides in aqueous solutions reported in Ref. [11], challenge the conventional notion that damage to the genome by ionizing radiation is mainly induced by OH radicals. The role of Dissociative Electron Transfer observed in real time is shown to possibly play a dominant role in inducing DNA damage. All these findings enhance the importance of theoretical works to help obtaining a comprehensive view of the physics and chemistry of the primary events subsequent to ionization.

So far, very few theoretical works have focused on the primary processes leading to DNA molecular fragmentation, and at the atomic level, geometry optimization and molecular dynamics are the two main tools of investigation. DNA damages have been investigated through quantum chemistry calculations. Works have mainly focused on singly ionized radicals of the DNA bases in the gas phase where geometries and energetics have been characterized, mostly for the purpose of assessing ionization energies, attacks by the OH radical, hydrogen abstraction, hole transfers, energetics of fragmentation pathways. Reviews on these calculations can be found in Refs. [12–17]. *Ab-initio* molecular dynamics is however the method of choice to follow chemical reactivity in time. This has in particular been applied to investigate the chemical reactivity of OH species on DNA bases in the gas phase by Mundy et al. [18,19]. We have developed a time-dependent method that allows to follow in time the primary events subsequent to the ionization process of a molecule within ionization tracks. However, before exposing the method it is convenient to describe some particular characteristics of the ion sources considered.

1.2 Swift heavy ions tracks

In the track of swift heavy ions, the energy released in the medium depends to a large extent on the velocity and the charge state of the ions. All along the track, the ion travels with a slowly growing Linear Energy Transfer (LET) until its energy reaches the weaker MeV and subMeV energy range. In the so-called high velocity regime (energy range > 1 MeV) the ion velocity is much larger than the velocity of the electrons of the target: the dominant phenomenon is the ionization of the target and consequently high energy secondary electrons tracks are produced. For example, 1 and 10 MeV/UMA Argon ions, with respectively 2600 and 1000 keV/mm LET, will eject several electrons from a water molecule with respectively about 150 and 500 eV mean kinetic energy upon multiple ionization [20]. When the energy of the ion is such that its velocity is of the order of magnitude of the velocity of the electrons of the target molecule, more complex phenomena occur, with the capture of electrons by the incident ion and the formation of an excited quasi-molecule (in particular with subsequent charge transfers between ion and target). As a consequence, the energy loss of the ion strongly increases, leading to a very excited and ionized target volume, until the ion stops. This particular region is called the Bragg peak, and this is the useful region in hadron therapy, since the velocity and the charge of the incident ion can be adjusted to target the adequate energy release for the depth of a tumour. Further complex physical phenomena occur in the Bragg peak. See, for instance, Refs. [21,22] for physical primary events and Refs. [23–25] for fragmentation events in ionized biomolecules for more details.

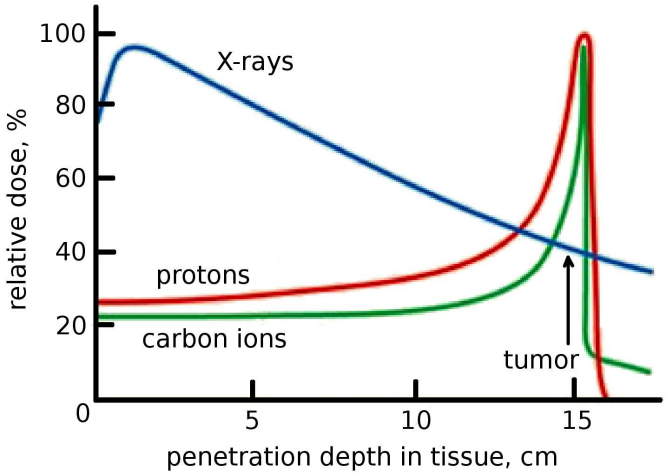


Figure 1.2: X-rays (photons) lose energy rapidly by ionization as they travel through the body. On the other hand, charged particles such as protons and carbon ions deposit most of their energy at a specific depth that depends on their energy (called the Bragg peak). This means that they can deliver a high radiation dose at a tumour site, while sparing the surrounding healthy tissue.

1.3 Working hypothesis

Our simulations are concerned and restricted to the first ≈ 100 fs after a double vertical ionization of molecules that accurr before the Bragg peak is reached. Both hypothesis about how the ionization is occurred can be argued as follows:

1. Multiple ionization: Multiple ionization and especially the double ionization arising from irradiation by swift heavy ions (about 10% of primary ionization events by ions), though

maybe less probable than other events resulting from the electronic cascading (for instance electronic excitations, electron attachments), may be systematically more damageable (and therefore more lethal to cells). This is supported by experiments that have been carried out in our group in the recent past [26–28]. Our present theoretical work thus focuses on demonstrating that molecular double ionization, either from water molecules of the solvent or from DNA molecules, may be in some cases responsible for the formation of an atomic oxygen as a direct consequence of the molecules Coulomb explosion. The chemical reactivity of the produced atomic oxygen with other radicals present in the medium will ultimately lead to chemical products that are harmful to DNA. In the case of pure liquid water, Gervais et al. [29] demonstrated that the multiple ionization hypothesis is indeed able to correctly interpret the experimental HO_2/O_2^- formation yield [30] for LET above 100 keV/μ as obtained with swift heavy ions. Furthermore, combining multi-scale theoretical modeling, we have shown that the production of HO_2 can only be achieved as a consequence of an initial double ionization of one H_2O molecule of the liquid [31].

2. Vertical ionization: In the LET range of swift heavy ions relevant to our work (typically 250 and 290 keV/mm for S^{16+} (77 MeV/UMA) and Ar^{18+}(70 MeV/UMA) ions), the experimental conditions (see, for instance, Refs. [30,32] for liquid water) are such that the path of ions is about 2.5 mm and the energy loss of the colliding ions is about 25% of the incident energy, which results in a 12% loss in velocity (velocity typically decreases from 0.38c to 0.34c). An associated short collision time of about 0.1 fs can thus be estimated. In that range of velocities, capture cross sections are negligible, so that the charge state of the colliding ions remains unchanged. Moreover, swift heavy ions projectiles have both velocity and mass high enough so that elastic collisions can be safely neglected, and consequently the energy of the ions upon collision is released into the target molecules (i.e. water molecules or biomolecules) through electronic excitation and ionization. For high collision energies, the ionization cross-section dominates all the other processes. For these ionization processes, the velocity of the incoming ion and of the extracted electrons are so high that both ions and electrons are swiftly evacuated from the collision zone: within 1 as (attosecond) subsequent to the collision and ionization events, both ions and electrons are not localized anymore in the vicinity of the ionized molecule. They thus do not interact further on. The vertical ionization of the target molecule is another consequence of these ultrashort time-scales of collision and interaction in ionizing irradiation by swift heavy ions, i.e. without any changes in the geometry or velocities of the target ionized molecule. In this context, the Coulomb explosion and fragmentation of the doubly ionized molecule must be considered in a highly ionized small volume (i.e. one single molecule ionized within the volume) within an ultra-short time-scale: the chemical processes responsible for the early stages of the fragmentation therefore take place within a few 100 of femtosecond after the collision and ionization events (femto-chemistry).

The ionization process by itself and the subsequent chemical processes occurring at the femtosecond time-scale are pure quantum events and thus should be modeled at the *ab-initio* level, i.e. taking into account the electronic wavefunction as governed by the Schrödinger equation (time-dependent or time-independent). To that end, we have developed a methodology based on Time-Dependent Density Functional Theory Molecular Dynamics simulations (TD-DFT MD) in which effective molecular orbitals are propagated in time [33, 34] and after we have used Born-Oppenheimer (BO) MD to exted the simulation to longer times. The key ingredient is the propagation of electronic effective molecular orbitals that are individually associated to each molecule of thesystem. With that methodology we can not only initiate multiple ionisation from any molecule of the sample (DNA molecule or any water molecule from the surrounding

bulk liquid) but we can also control the molecular orbital from which this multiple ionisation proceeds. All our calculations are performed with the electronic representation at the Density Functional Theory (DFT) level as a good compromise between accuracy, size of the simulated system and affordable calculation time (CPU time).

The molecules aim of our study have been uracil, pirimidine and small water cluster. The first two have been chosen regarding their relevance as basic componets of RNA. This choice is also supported by the rich experimental data available [24,25,35–42] based on simplicity and high stability in gas phase (which prevents, e.g., its thermal degradation) of both molecules. On the other hand, ionization of water clusters is relevant to radiation damage studies because water is the natural environment of biomolecules. Regarding this issue, uracil have been also examinated in the bulk liquid phase at finite temperature, taking into account the solute biomolecule and the surrounding liquid water molecules at the DFT level, all molecules being explicitly described. These simulation have allow us to give insights to the direct versus indirect effects involved in double ionizations.

1.4 Structure of this thesis

This dissertation is organized as follows:

The second part presents the foundations of the methodology organized in three chapters. First one is devoted to expose the main approaches of the Molecular Dynamics (MD) performed in this thesis, namely: Born-Oppenhemier MD, Ehrenfest MD and Car-Parrinello MD. In the following two chapters, we will explain the implementation of such approximations in DFT and TD-DFT framework and also we will We also

The results will be presented in the third part of the thesis. The chapters are based on articles published (or submitted) in international journals of chemistry and physics and are classified according to the system studied. Thus, Chapters 5 and 6 deal with double ionizations of uracil in the gas and liquid phase respectively; Chapter 7 with double ionization of small water clusters and, finally, Chapter 8 shows the results for double ionizations of pyrimidine. We have extracted the most relevant conclusions in the fourth part of the thesis. Some appendices are also included at the end to give details on mathematical deductions, fragmentation channels and dissociation barriers.

Part II

Methodology

Chapter 2

Ab-initio Molecular Dynamics

2.1 Introduction

Molecular Dynamics (MD) simulations has had a long history [43–47] and has evolved into an important and widely used theoretical tool that allow to model different types of systems, including gases, liquids, solids, surfaces and clusters [48–51]. It aims at simulating both nuclei and electrons all together, although, due to the complex description of their interactions, approaches have to be considered. The methods can be then classified in several groups: Classical MD simulations based on Newtons second law [52], Force Fields MD [53] in which the nuclei evolve according to an empirical potential designed as close as possible to the real system, Quantum-Classical Molecular Dynamics (QCMD) [54–56] (or *Ab-initio* MD methods) for which quantum description is used for electrons while the nuclei are kept as a classical particles, and Quantum MD simulations that aim to a more complex quatum description of electrons and/or nuclei [57, 58]. The choice of one method over the others mainly concerns the information that we want to extract from the simulation, though in a practical sense, factors like computational recourses are also limiting.

In this Chapter we will summarize the theoretical concepts behind the MD techniques used in this thesis. All of them belong to the group of QCMD (nuclei treated classically/electrons using quantum mechanics). There are many of such methods that can be at the same time classified attending to the approximations made in the description of the electronic movement [59]. The ones used here are included in the so-called *on-the-fly* methods for which the total calculation of the Potential Energy Surface (PES) is rejected. Instead, the surface is reduced only to those nuclear configurations given by the classical trajectory. Namely, we will describe the essentials of Born-Oppenheimer, Ehrenfest and Car-Parrinello approximations first using a wavefunction-based formulation [60, 61], and afterwards, we will go through their implementation in the Density Functional Theory (DFT) and Time-Dependent DFT (TD-DFT).

2.2 Schrödinger Equation and Molecular Hamiltonian

We are interested in describing systems formed by electrons and nuclei. According to wave mechanics proposed by E. Schrödinger [62], the physical stage of such systems can be mathematically expressed by a wavefunction $|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle$ where $\mathbf{r}$ and $\mathbf{R}$ represent electronic and nuclear degrees of freedom. This function contains all the information that is possible to know

about the system and its time variation is postulated as follows:

$$i\hbar\frac{\partial}{\partial t}|\Phi(\mathbf{r},\mathbf{R},t)\rangle = \hat{H}|\Phi(\mathbf{r},\mathbf{R},t)\rangle \tag{2.1}$$

The Eq.(2.1) is known as *Time-dependent Schrödinger Equation* (TDSE), where the non-relativistic Hamiltonian operator $\hat{H}$, given a system of N nuclei and n electrons, is defined by:

$$\hat{H} = -\sum_{\alpha=1}^{N}\frac{\hbar^2}{2M_\alpha}\nabla_\alpha^2 - \sum_{i=1}^{n}\frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_{\alpha=1}^{N}\sum_{i=1}^{n}\frac{Z_\alpha q_e^2}{r_{\alpha i}} + \sum_{\alpha=1}^{N}\sum_{\beta>\alpha}^{N}\frac{Z_\alpha Z_\beta q_e^2}{r_{\alpha\beta}} + \sum_{i=1}^{n}\sum_{j>i}^{n}\frac{q_e^2}{r_{ij}} \tag{2.2}$$

where $\nabla^2 = \frac{\partial^2}{\partial x^2}\mathbf{i} + \frac{\partial^2}{\partial y^2}\mathbf{j} + \frac{\partial^2}{\partial z^2}\mathbf{k}$, subindex α and β denote nuclei, and i and j electrons.

The first term of Eq.(2.2) is the kinetic nuclear energy operator while the second term represents the kinetic electron energy operator. The third term is the potential nuclear-electron attraction energy operator being $r_{\alpha i}$ the distance between nucleus α and electron i. The fourth term represents the nuclear-nuclear repulsion operator where $r_{\alpha\beta}$ is the distance between nuclei α and β. Finally, last term is the electron-electron repulsion operator being r_{ij} the distance between electron i and j.

Analogous expressions for Eq.(2.2) will be used throughout this text:

$$\begin{aligned}\hat{H} &= -\sum_{\alpha=1}^{N}\frac{\hbar^2}{2M_\alpha}\nabla_\alpha^2 - \sum_{i=1}^{n}\frac{\hbar^2}{2m_e}\nabla_i^2 + \hat{V}_{n-e}(\mathbf{r},\mathbf{R}) = \hat{T}_N(\mathbf{R}) + \hat{T}_e(\mathbf{r}) + \hat{V}_{n-e}(\mathbf{r},\mathbf{R}) = \\ &= \hat{T}_N(\mathbf{R}) + \hat{T}_e(\mathbf{r}) + \hat{V}_{eN}(\mathbf{r},\mathbf{R}) + \hat{V}_{NN}(\mathbf{R}) + \hat{V}_{ee}(\mathbf{r}) = \hat{T}_N(\mathbf{R}) + \hat{H}_e(\mathbf{r},\mathbf{R})\end{aligned} \tag{2.3}$$

where $\hat{T}_N(\mathbf{R})$ is the kinetic nuclear operator, $\hat{T}_e(\mathbf{r})$ is the kinetic electronic operator and $\hat{H}_e(\mathbf{r},\mathbf{R})$ is the electronic Hamiltonian. If there is not any time-dependent external field applied, the molecular Hamiltonian is explicitly time-independent. A *particular* solution of the TDSE can be found separating the electronic and nuclear variables from time:

$$|\Phi(\mathbf{r},\mathbf{R},t)\rangle = |\psi(\mathbf{r},\mathbf{R})\rangle|f(t)\rangle \tag{2.4}$$

Replacing Eq.(2.4) in Eq.(2.1) we get:

$$-\frac{\hbar}{i}\frac{1}{|f(t)\rangle}\frac{\partial|f(t)\rangle}{\partial t} = \left[\sum_{\alpha=1}^{N}\frac{\hbar^2}{2M_\alpha}\nabla_\alpha^2 + \hat{H}_e(\mathbf{r},\mathbf{R})\right]|\psi(\mathbf{r},\mathbf{R})\rangle \tag{2.5}$$

Eq.(2.5) is separable into two independent differential equations. The first one is a time-dependent equation:

$$\frac{d|f(t)\rangle}{|f(t)\rangle} = -\frac{iE}{\hbar}dt \tag{2.6}$$

While the second one is a spatial-dependent eigenvalue equation:

$$[\hat{T}_N(\mathbf{R}) + \hat{H}_e(\mathbf{r},\mathbf{R})]|\psi(\mathbf{r},\mathbf{R})\rangle = E|\psi(\mathbf{r},\mathbf{R})\rangle \tag{2.7}$$

that is called the *Time-independent Schrödinger Equation* (TISE). The constant E is a real number (since $\hat{H}$ is an hermitian operator) that corresponds to the energy of the state represented by the eingenfunction $|\psi(\mathbf{r},\mathbf{R})\rangle$. Notice that the number of solutions of the TISE can vary from

a discrete number of them for bounded systems to a continuous range in unbounded potentials. Differential equations like Eq.(2.6) have known solutions:

$$|f(t)\rangle = A\exp(\frac{-i}{\hbar}Et) \tag{2.8}$$

Then, the total wavefunction solutions of the TISE can be written as:

$$|\Phi(\mathbf{r},\mathbf{R},t)\rangle = |\psi(\mathbf{r},\mathbf{R})\rangle\exp(\frac{-i}{\hbar}Et) \tag{2.9}$$

where the amplitude A is included in $|\psi(\mathbf{r},\mathbf{R})\rangle$. Eq.(2.9) represent states in which the system is moving but with constant energy and also constant probability density since:

$$\langle\Phi(\mathbf{r},\mathbf{R},t)|\Phi(\mathbf{r},\mathbf{R},t)\rangle = |\psi(\mathbf{r},\mathbf{R})|^2 \tag{2.10}$$

It is worth noting that in both TDSE and TISE there is a time-dependence of the total wavefunction. In the later, however, the dependence is given as a wave phase that makes the energy and also the probability density time-independent.

A huge number of chemical problems have been successfully studied relying on these stationary states. Specially, the state that corresponds to the smallest energy in Eq.(2.7) (*ground state*) plays an important role for many chemical and physical process. On the contrary, if we were interested in cases in which the Hamiltonian depends on time (e.g. a laser source); a more complex wavefunction has to be used to really represent the time dependence. Indeed, we must be aware that Eq.(2.9) is only a possible solution of the TDSE since it can be demonstrated [63] that linear combinations of different stationary solutions:

$$|\Phi(\mathbf{r},\mathbf{R},t)\rangle = \sum_{n=0}^{\infty} a_n|\psi_n(\mathbf{r},\mathbf{R})\rangle\exp(\frac{-i}{\hbar}E_nt) \tag{2.11}$$

are also solution of the TDSE due to the linearity of the proper time-dependent equation. If the spectra is continuous summation can be changed into an integral:

$$|\Phi(\mathbf{r},\mathbf{R},t)\rangle = \int_o^{\infty} a(E)|\psi_E(\mathbf{r},\mathbf{R})\rangle\exp(\frac{-i}{\hbar}Et)dE \tag{2.12}$$

we obtain a *wavepacket*, that represents one way to include the nontrivial time dependence of the wavefunction.

The complexity of the total wavefunction, $|\Phi(\mathbf{r},\mathbf{R},t)\rangle$, for big molecular systems far from diatomic molecules becomes unaffordable for many practical pourposes. Mainly, a pure wavefunction involves degrees of freedom of two different natures (nuclear and electrons) and their variation on time is really hard to capture properly with theoretical models. Approximations are arisen to overcome these problems and basically aim to decrease the complexity of the wavefunction by imposing simpler expressions.

2.3 Hellmann-Feynman theorem

In QCMD, the major problem is concentrated on the accurately calculation of the PES and also on the description of the feedback between electrons and nuclei. In the later, the derivatives of the PES are really important since nuclear forces are extrated from them. Unfortunately, analytical derivations of such surfaces are highly demanding in terms of CPU time. Nowadays, most of the procedures implemented in the codes lay on the Hellmann-Feynman (HF) theorem that supposes a much easier way to calculate nuclear forces.

2.3.1 Hellmann-Feynman for time-independent wavefunction

Let's consider the quantum definition of force as the spectation value of the time derivative of the momentum operator [64]:

$$\begin{aligned}\mathbf{F}_\lambda &= \frac{\hbar}{i}\frac{d}{dt}\langle\frac{\partial}{\partial\lambda}\rangle = \frac{\hbar}{i}\frac{d}{dt}\langle\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial}{\partial\lambda}|\Phi(\mathbf{r},\mathbf{R},t)\rangle = \\ &= \frac{\hbar}{i}\langle\frac{\partial}{\partial t}\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial}{\partial\lambda}|\Phi(\mathbf{r},\mathbf{R},t)\rangle + \frac{\hbar}{i}\langle\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial}{\partial\lambda}|\frac{\partial}{\partial t}\Phi(\mathbf{r},\mathbf{R},t)\rangle \end{aligned} \tag{2.13}$$

where λ can be any electronic or nuclear degree of freedom[1]

If the wavefunction is stationary, i.e. verifies the Eq.(2.9), then:

$$\mathbf{F}_\lambda = \langle\psi(\mathbf{r},\mathbf{R})|E\frac{\partial}{\partial\lambda}|\psi(\mathbf{r},\mathbf{R})\rangle - \langle\psi(\mathbf{r},\mathbf{R})|\frac{\partial}{\partial\lambda}E|\psi(\mathbf{r},\mathbf{R})\rangle \tag{2.14}$$

Considering the hermeticity of the Hamiltonian:

$$\mathbf{F}_\lambda = \langle\psi(\mathbf{r},\mathbf{R})|\hat{H}\frac{\partial}{\partial\lambda}|\psi(\mathbf{r},\mathbf{R})\rangle - \langle\psi(\mathbf{r},\mathbf{R})|\frac{\partial}{\partial\lambda}\hat{H}|\psi(\mathbf{r},\mathbf{R})\rangle = \langle[H,\frac{\partial}{\partial\lambda}]\rangle \tag{2.15}$$

Some useful equivalences taken from Ref. [65]:

$$\langle[H,\frac{\partial}{\partial\lambda}]\rangle = -\langle[\frac{\partial}{\partial\lambda},H]\rangle = -\frac{i}{\hbar}\langle[p,H]\rangle = -\langle\psi(\mathbf{r},\mathbf{R})|\frac{\partial}{\partial\lambda}\hat{H}|\psi(\mathbf{r},\mathbf{R})\rangle \tag{2.16}$$

Therefore, particularizing Eq.(2.14) to a nuclear position ($\lambda = \alpha$), we get:

$$\mathbf{F}_\alpha = -\langle\Psi(\mathbf{r},\mathbf{R})|\nabla_\alpha\hat{H}|\Psi(\mathbf{r},\mathbf{R})\rangle \tag{2.17}$$

that represents the force acting on a particular nuclei. On the other hand, the energy of the system is given by the TISE (2.7) expressed in a mean value form:

$$E = \langle\Psi(\mathbf{r},\mathbf{R})|\hat{H}|\Psi(\mathbf{r},\mathbf{R})\rangle \tag{2.18}$$

The variation of such expected value with any of the nuclear coordinates, $\mathbf{R}_\alpha$, is defined by[2]:

$$\begin{aligned}\nabla_\alpha E = \nabla_\alpha\langle\Psi(\mathbf{r},\mathbf{R})|\hat{H}|\Psi(\mathbf{r},\mathbf{R})\rangle &= \langle\nabla_\alpha\Psi(\mathbf{r},\mathbf{R})|\hat{H}|\Psi(\mathbf{r},\mathbf{R})\rangle + \langle\Psi(\mathbf{r},\mathbf{R})|\nabla_\alpha\hat{H}|\Psi(\mathbf{r},\mathbf{R})\rangle + \\ &+ \langle\Psi(\mathbf{r},\mathbf{R})|\hat{H}|\nabla_\alpha\Psi(\mathbf{r},\mathbf{R})\rangle = E\langle\nabla_\alpha\Psi(\mathbf{r},\mathbf{R})|\Psi(\mathbf{r},\mathbf{R})\rangle + \langle\Psi(\mathbf{r},\mathbf{R})|\nabla_\alpha\hat{H}|\Psi(\mathbf{r},\mathbf{R})\rangle + \\ &+ E\langle\Psi(\mathbf{r},\mathbf{R})|\nabla_\alpha\Psi(\mathbf{r},\mathbf{R})\rangle \end{aligned} \tag{2.19}$$

Since the electronic wavefunction is normalized:

$$\langle\Psi(\mathbf{r},\mathbf{R})|\Psi(\mathbf{r},\mathbf{R})\rangle = 1 \Rightarrow \nabla_\alpha\langle\Psi(\mathbf{r},\mathbf{R})|\Psi(\mathbf{r},\mathbf{R})\rangle = 0 \tag{2.20}$$

Eq.(2.19) is reduced to the following expression:

$$\nabla_\alpha E = \langle\Psi(\mathbf{r},\mathbf{R})|\nabla_\alpha\hat{H}|\Psi(\mathbf{r},\mathbf{R})\rangle \tag{2.21}$$

Combining both Eq.(2.17) and Eq.(2.21), finally we obtain the HF theorem:

$$\mathbf{F}_\alpha = -\langle\Psi(\mathbf{r},\mathbf{R})|\nabla_\alpha\hat{H}|\Psi(\mathbf{r},\mathbf{R})\rangle = -\nabla_\alpha E \tag{2.22}$$

[1]Notation: the symbol | is completely optional, i.e. $\langle\phi|A|\psi\rangle = \langle\phi|A\psi\rangle$

[2]Notation: $\langle\nabla_\alpha\Psi(\mathbf{r},\mathbf{R})|\hat{H}|\Psi(\mathbf{r},\mathbf{R})\rangle = \int\nabla_\alpha\Psi^*(\mathbf{r},\mathbf{R})\hat{H}\Psi(\mathbf{r},\mathbf{R})d\tau$

that was proven by many authors (Paul Güttinger [66], Wolfgang Pauli [67], Hans Hellmann [68] and Richard Feynman [69]) independently. As we said at the beginning of the section, the practical importance of the HF theorem is that energy derivatives are difficult to compute numerically, whereas the expression in Eq.(2.17) can be computed efficiently. For instance, in the molecular problem, only two components of the Hamiltonian contribute to the required derivative −the electron-nucleus and nucleus-nucleus terms:

$$\frac{\partial \hat{H}}{\partial \mathbf{R}_\gamma} = \frac{\partial}{\partial \mathbf{R}_\gamma}\left[-\sum_{\alpha=1}^{N}\sum_{i=1}^{n}\frac{Z_\alpha q_e^2}{r_{\alpha i}} + \sum_{\alpha=1}^{N}\sum_{\beta>\alpha}^{N}\frac{Z_\alpha Z_\beta q_e^2}{r_{\alpha\beta}}\right] = \\ = Z_\gamma q_e^2 \sum_{i=1}^{n}\frac{\mathbf{R}_\gamma - \mathbf{r}_i}{|\mathbf{R}_\gamma - \mathbf{r}_i|^3} - Z_\gamma q_e^2 \sum_{\alpha\neq\gamma}^{N} Z_\alpha \frac{\mathbf{R}_\gamma - \mathbf{r}_i}{|\mathbf{R}_\gamma - \mathbf{R}_\alpha|^3} \tag{2.23}$$

where $\mathbf{R}_\gamma$ and $\mathbf{r}_i$ are nuclear and electronic positions. Nevertheless, it was recognized by early computational work using atom-centered basis functions that the HF force gave manifestly wrong results [70]. In fact, whenever a variational wavefunction is used:

$$E\langle\Psi(\mathbf{r},\mathbf{R})|\Psi(\mathbf{r},\mathbf{R})\rangle = \langle\Psi(\mathbf{r},\mathbf{R})|\hat{H}|\Psi(\mathbf{r},\mathbf{R})\rangle \tag{2.24}$$

where E and $|\Psi(\mathbf{r},\mathbf{R})\rangle$ now represent the variational energy and the variational (unnormalized) wavefunction respectively. The Eq.(2.19) is transformed in the following expression:

$$\nabla_\alpha E\langle\Psi(\mathbf{r},\mathbf{R})|\Psi(\mathbf{r},\mathbf{R})\rangle = \langle\Psi(\mathbf{r},\mathbf{R})|\nabla_\alpha\hat{H}|\Psi(\mathbf{r},\mathbf{R})\rangle + \\ + \langle\nabla_\alpha\Psi(\mathbf{r},\mathbf{R})|\hat{H}-E|\Psi(\mathbf{r},\mathbf{R})\rangle + \langle\Psi(\mathbf{r},\mathbf{R})|\hat{H}-E|\nabla_\alpha\Psi(\mathbf{r},\mathbf{R})\rangle \tag{2.25}$$

It can be shown that Eq.(2.25) is reduced to the HF theorem (2.22) for exact variational wavefunctions (e.g. Hartree-Fock) expanded in a infinite basis set [71]. If a finite localized basis sets are used however, second and third term of Eq.(2.25) (known as a Pulay forces [70]) have to be evaluated explicitly. On the contrary, using originless basis functions, such as plane waves, Pulay forces vanishes exactly since they do not depend on $\mathbf{R}_\alpha$. That applies of course to all *ab-initio* molecular dynamics schemes using that particular basis set.

2.3.2 Hellmann-Feynman for time-dependent wavefunction

We now address the issue of connecting the rigorous force definition to the HF force for a normalized time-dependent wavefunction, $|\Phi(\mathbf{r},\mathbf{R},t)\rangle$, that satisfies the TDSE (2.1):

$$\mathbf{F}_\lambda = \frac{\hbar}{i}\frac{d}{dt}\langle\frac{\partial}{\partial\lambda}\rangle = \langle\hat{H}\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial}{\partial\lambda}|\Phi(\mathbf{r},\mathbf{R},t)\rangle + \\ + \frac{\hbar}{i}\langle\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial}{\partial t}\frac{\partial}{\partial\lambda}|\Phi(\mathbf{r},\mathbf{R},t)\rangle - \langle\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial}{\partial\lambda}\hat{H}|\Phi(\mathbf{r},\mathbf{R},t)\rangle \tag{2.26}$$

that is nothing but the Ehrenfest theorem detailed in appendix A.1:

$$\mathbf{F}_\lambda = \frac{\hbar}{i}\left[\langle\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial}{\partial t}\frac{\partial}{\partial\lambda}|\Phi(\mathbf{r},\mathbf{R},t)\rangle + \frac{i}{\hbar} < [\hat{H},\frac{\partial}{\partial\lambda}] >\right] \tag{2.27}$$

in which the first term in brackets is equal to zero. Therefore:

$$\mathbf{F}_\lambda = < [\hat{H},\frac{\partial}{\partial\lambda}] > = -\frac{i}{\hbar}\langle[p,H]\rangle = -\langle\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial}{\partial\lambda}\hat{H}|\Phi(\mathbf{r},\mathbf{R},t)\rangle \tag{2.28}$$

that for a particular nuclear position ($\lambda = \alpha$):

$$\mathbf{F}_\alpha = -\langle\Phi(\mathbf{r},\mathbf{R},t)|\nabla_\alpha\hat{H}|\Phi(\mathbf{r},\mathbf{R},t)\rangle \tag{2.29}$$

Let's analyze the derivative of the expectation value of the energy to any of the degrees of freedom, λ:

$$\begin{aligned}
\frac{\partial}{\partial\lambda}\langle\Phi(\mathbf{r},\mathbf{R},t)|\hat{H}|\Phi(\mathbf{r},\mathbf{R},t)\rangle = \\
&= \langle\frac{\partial}{\partial\lambda}\Phi(\mathbf{r},\mathbf{R},t)|\hat{H}|\Phi(\mathbf{r},\mathbf{R},t)\rangle + \langle\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial}{\partial\lambda}\hat{H}|\Phi(\mathbf{r},\mathbf{R},t)\rangle + \\
&+ \langle\Phi(\mathbf{r},\mathbf{R},t)|\hat{H}|\frac{\partial}{\partial\lambda}\Phi(\mathbf{r},\mathbf{R},t)\rangle = i\hbar\langle\frac{\partial}{\partial\lambda}\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial}{\partial t}\Phi(\mathbf{r},\mathbf{R},t)\rangle + \\
&+ \langle\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial}{\partial\lambda}\hat{H}|\Phi(\mathbf{r},\mathbf{R},t)\rangle - i\hbar\langle\frac{\partial}{\partial t}\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial}{\partial\lambda}\Phi(\mathbf{r},\mathbf{R},t)\rangle = \\
&= i\hbar\frac{d\lambda}{dt}\langle\frac{\partial}{\partial\lambda}\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial}{\partial\lambda}\Phi(\mathbf{r},\mathbf{R},t)\rangle + \langle\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial}{\partial\lambda}\hat{H}|\Phi(\mathbf{r},\mathbf{R},t)\rangle - \\
&- i\hbar\frac{d\lambda}{dt}\langle\frac{\partial}{\partial\lambda}\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial}{\partial\lambda}\Phi(\mathbf{r},\mathbf{R},t)\rangle = \langle\Phi(\mathbf{r},\mathbf{R},t)|\frac{\partial\hat{H}}{\partial\lambda}|\Phi(\mathbf{r},\mathbf{R},t)\rangle
\end{aligned} \tag{2.30}$$

where the chain rule has been applied.

From Eq.(2.29) and Eq.(2.30), finally we get the HF for time-dependent wavefunctions:

$$\mathbf{F}_\alpha = -\langle\Phi(\mathbf{r},\mathbf{R},t)|\nabla_\alpha\hat{H}|\Phi(\mathbf{r},\mathbf{R},t)\rangle = -\nabla_\alpha E \tag{2.31}$$

For a variational wavefunction a more complex expression of forces, similar to Eq.(2.25), can be also derived. The resulting Pulay forces are zero if the basis functions do not depend on the nuclear coordinates (see Ref. [72]).

In conclusion, we have demonstrated that forces acting on a nuclei can be obtained by deriving the mean value of the electronic energy to the selected nuclear coordinate. This statement is valid for both time-dependent and time-independent wavefunction whenever a delocalized or complete basis sets is used.

2.4 The Born-Oppeheimer approximation

Regarding the TISE, solutions to this equation cannot be suggested as a simple separation of electronic and nuclear variables, $|\Psi_e(\mathbf{r})\rangle|\chi_N(\mathbf{R})\rangle$, since the molecular Hamiltonian operator has a crossed term, $\hat{V}_{eN}(\mathbf{r},\mathbf{R})$, that depends on the two sets of coordinates. Nevertheless, the total spatial function might be expressed as a product of a electronic wavefunction and a nuclear wavefunction:

$$|\psi(\mathbf{r},\mathbf{R})\rangle = |\Psi_k(\mathbf{r};\mathbf{R})\rangle|\chi_N(\mathbf{R})\rangle \tag{2.32}$$

Where electronic wavefunction has a parametric dependency on nuclear positions that is pointed out by the semicolon. Many times this product is referred as the Born-Oppenheimer (BO) approximation, but, wavefunctions like Eq.(2.32) have exact solution of the TISE (2.7), i.e.:

$$\hat{H}|\Psi_k(\mathbf{r};\mathbf{R})\rangle|\chi_N(\mathbf{R})\rangle = E|\Psi_k(\mathbf{r};\mathbf{R})\rangle|\chi_N(\mathbf{R})\rangle \tag{2.33}$$

Or equivalently:

$$\begin{aligned}
\{\hat{T}_N(\mathbf{R}) + \hat{T}_e(\mathbf{r}) + \hat{V}_{eN}(\mathbf{r},\mathbf{R}) + \hat{V}_{NN}(\mathbf{R}) + \hat{V}_{ee}(\mathbf{r})\}|\Psi_k(\mathbf{r};\mathbf{R})\rangle|\chi_N(\mathbf{R})\rangle = \\
= E|\Psi_k(\mathbf{r};\mathbf{R})\rangle|\chi_N(\mathbf{R})\rangle
\end{aligned} \tag{2.34}$$

By the chain rule:

$$\begin{aligned}&\hat{T}_N(\mathbf{R})|\Psi_k(\mathbf{r};\mathbf{R})\rangle|\chi_N(\mathbf{R})\rangle = \\ &= -\sum_{\alpha=1}^{N}\frac{\hbar^2}{2M_\alpha}|\Psi_k(\mathbf{r};\mathbf{R})\rangle\nabla_\alpha^2|\chi_N(\mathbf{R})\rangle - \sum_{\alpha=1}^{N}\frac{\hbar^2}{M_\alpha}\nabla_\alpha|\Psi_k(\mathbf{r};\mathbf{R})\rangle\nabla_\alpha|\chi_N(\mathbf{R})\rangle - \\ &\quad -\sum_{\alpha=1}^{N}\frac{\hbar^2}{2M_\alpha}|\chi_N(\mathbf{R})\rangle\nabla_\alpha^2|\Psi_k(\mathbf{r};\mathbf{R})\rangle \end{aligned} \tag{2.35}$$

Since $\hat{T}_e(\mathbf{r})$ contains no $\mathbf{R}$ dependence:

$$\hat{T}_e(\mathbf{r})|\chi_k(\mathbf{r};\mathbf{R})\rangle|\Psi_N(\mathbf{R})\rangle = |\chi_N(\mathbf{R})\rangle\hat{T}_e(\mathbf{r})|\Psi_k(\mathbf{r};\mathbf{R})\rangle \tag{2.36}$$

and considering the parametric dependence:

$$\hat{V}_{eN}|\Psi_k(\mathbf{r};\mathbf{R})\rangle|\chi_N(\mathbf{R})\rangle = |\chi_N(\mathbf{R})\rangle\hat{V}_{eN}|\Psi_k(\mathbf{r};\mathbf{R})\rangle \tag{2.37}$$

If we rearrange terms:

$$\begin{aligned}&-\sum_{\alpha=1}^{N}\frac{\hbar^2}{2M_\alpha}|\Psi_k(\mathbf{r};\mathbf{R})\rangle\nabla_\alpha^2|\chi_N(\mathbf{R})\rangle - \sum_{\alpha=1}^{N}\frac{\hbar^2}{M_\alpha}\nabla_\alpha|\Psi_k(\mathbf{r};\mathbf{R})\rangle\nabla_\alpha|\chi_N(\mathbf{R})\rangle - \\ &-\sum_{\alpha=1}^{N}\frac{\hbar^2}{2M_\alpha}|\chi_N(\mathbf{R})\rangle\nabla_\alpha^2|\Psi_k(\mathbf{r};\mathbf{R})\rangle + |\chi_N(\mathbf{R})\rangle\hat{T}_e|\Psi_k(\mathbf{r};\mathbf{R})\rangle + \\ &+|\chi_N(\mathbf{R})\rangle\hat{V}_{eN}(\mathbf{R})|\Psi_k(\mathbf{r};\mathbf{R})\rangle + |\Psi_k(\mathbf{r};\mathbf{R})\rangle\hat{V}_{NN}(\mathbf{R})|\chi_N(\mathbf{R})\rangle + \\ &+|\chi_N(\mathbf{R})\rangle\hat{V}_{ee}(\mathbf{R})|\Psi_k(\mathbf{r};\mathbf{R})\rangle = E|\Psi_k(\mathbf{r};\mathbf{R})\rangle|\chi_N(\mathbf{R})\rangle \end{aligned} \tag{2.38}$$

Until this point we have been using equalities meaning that the product of Eq.(2.32) is exact solution. The Born-Oppenheimer approximation can be introduced here by estimating the third kinetic term of Eq.(2.38). The basis of the approximation is held by the fact that nuclei are much heavier than electrons. Keeping this mass difference, we can see that a typical contribution of this term, in atomic units, has the form $1/(2M_\alpha)\nabla_\alpha^2|\Psi_k(\mathbf{r};\mathbf{R})\rangle$ [73]. The value of $\nabla_\alpha|\Psi_k(\mathbf{r};\mathbf{R})\rangle$ is of the same order as $\nabla_i|\Psi_k(\mathbf{r};\mathbf{R})\rangle$ since the derivatives operate over approximately the same dimensions. The latter is $|\Psi_k(\mathbf{r};\mathbf{R})\rangle p_e$, with p_e the momentum of an electron. Therefore $1/(2M_\alpha)\nabla_A^2|\psi_k(\mathbf{r};\mathbf{R})\rangle \approx p_e^2/(2M_\alpha) = (m_e/M_\alpha)T_e$. Since $m/M_\alpha \sim 1/1000$, the third can be dropped, and also the second since it is even much smaller. Ignoring both, we get:

$$\begin{aligned}&|\Psi_k(\mathbf{r};\mathbf{R})\rangle\hat{T}_N(\mathbf{R})|\chi_N(\mathbf{R})\rangle + |\chi_N(\mathbf{R})\rangle\hat{T}_e(\mathbf{r})|\Psi_k(\mathbf{r};\mathbf{R})\rangle + |\Psi_k(\mathbf{r};\mathbf{R})\rangle\hat{V}_{eN}(\mathbf{r},\mathbf{R})\rangle|\chi_N(\mathbf{R})\rangle + \\ &+|\Psi_k(\mathbf{r};\mathbf{R})\rangle\hat{V}_{NN}(\mathbf{R})|\chi_N(\mathbf{R})\rangle + |\chi_N(\mathbf{R})\rangle\hat{V}_{ee}(\mathbf{r})|\Psi_k(\mathbf{r};\mathbf{R})\rangle \approx E|\Psi_k(\mathbf{r};\mathbf{R})\rangle|\chi_N(\mathbf{R})\rangle \end{aligned} \tag{2.39}$$

On the other hand, due to the mass difference, the nuclei move very slowly in comparison with electrons and they can be considered as nearly fixed with respect to electron movement. The TISE (2.7) that represents such system is:

$$\hat{H}_e(\mathbf{r},\mathbf{R})|\Psi_k(\mathbf{r};\mathbf{R})\rangle = E_k(\mathbf{R})|\Psi_k(\mathbf{r};\mathbf{R})\rangle \tag{2.40}$$

where the constant nuclear-nuclear interaction potential V_{NN} has been included in $E_k(\mathbf{R})$. The electronic level to which corresponds the electronic wavefunction is expressed by the subindex "k". The number of electronic states is not unique for every molecule, in fact k can run from a discrete number to infinite. Linear combinations of these states can be also proposed as a solution

of the TISE leading to coupling between states. If we constrain the electrons to be in only one of them (*adiabatic approximation*) the problem of coupling between them is eliminated since all coupling terms go to zero. In this respect, M. Born and R. Oppenheimer in 1927 [74] studied, by a perturbative analysis of a time-independent Hamiltonian, the influence of the nuclei on the electronic wavefunction in the mass ratio $k = (m_e/M)^{1/4}$, being M the average mass weight of the molecule. They concluded that electrons do not undergo transitions between states under *appropriate* conditions. Even though, there are many situations where the adiabatic separation breaks down [65, 75], they referred basically to those cases in which the scale of the electronic energy and the nuclear vibrational motion are not comparable.

Thus, regarding the Eq.(2.40), Eq.(2.39) can be re-write as:

$$\left(\hat{T}_N + \langle\Psi_k(\mathbf{r};\mathbf{R})|\hat{H}_e(\mathbf{r},\mathbf{R})|\Psi_k(\mathbf{r};\mathbf{R})\rangle\right)|\chi_N(\mathbf{R})\rangle \approx E|\chi_N(\mathbf{R})\rangle \tag{2.41}$$

We have arrived to a singular scheme in which the nuclei are moving on a PES generated by the electrons. Due to the mass difference, the molecular problem is translated to a new picture where the nuclei follow their dynamics, and the electrons instantaneously adjust their wavefunction according to the nuclear wavefunction. We also guess that electrons will remain always in the same stationary state of the electronic Hamiltonian making the transitions between states only possible through the coupling with an external electromagnetic field.

2.5 Born-Oppenheimer ansatz and electronic couplings

Let us suppose the solutions of Eq.(2.40) are known and assume that the spectrum of $\hat{H}_e(\mathbf{r},\mathbf{R})$ is discrete. If the eigenfunctions are orthonormalized:

$$\langle\Psi_k(\mathbf{r};\mathbf{R})|\Psi_l(\mathbf{r};\mathbf{R})\rangle = \delta_{k,l} \tag{2.42}$$

the total time-dependent wavefunction $|\Phi(\mathbf{r},\mathbf{R};t)\rangle$ can be expanded in terms of the eigenfunctions of $\hat{H}_e$ since these form a complete set:

$$|\Phi(\mathbf{r},\mathbf{R},t)\rangle = \sum_{l=0}^{\infty}|\Psi_l(\mathbf{r};\mathbf{R})\rangle|\chi_l(\mathbf{R},t)\rangle \tag{2.43}$$

This product constitutes the so-called BornOppenheimer ansatz [65, 76, 77] and represents the simplest way to include the time variable into the the total wavefunction. Insertion of this ansatz into the time-dependent Schrödinger Eq.(2.1) followed by multiplication from the left by $\langle\Psi_k(\mathbf{r};\mathbf{R})|$ leads to a set of coupled differential equations (see appendix A.2 for details):

$$i\hbar\frac{\partial|\chi_k(\mathbf{R},t)\rangle}{\partial t} = [\hat{T}_N(\mathbf{R}) + E_k(\mathbf{R})]|\chi_k(\mathbf{R},t)\rangle + \sum_{l=0}^{\infty}C_{kl}|\chi_l(\mathbf{R},t)\rangle \qquad k = 0,1,2,...,\infty \tag{2.44}$$

where the exact coupling operator C_{kl} is defined as:

$$C_{kl} = \langle\Psi_k(\mathbf{r};\mathbf{R})|\hat{T}_N(\mathbf{R})|\Psi_l(\mathbf{r};\mathbf{R})\rangle - \sum_{\alpha=1}^{N}\frac{\hbar^2}{M_\alpha}\{\langle\Psi_k(\mathbf{r};\mathbf{R})|\nabla_\alpha|\Psi_l(\mathbf{r};\mathbf{R})\rangle\}\nabla_\alpha \tag{2.45}$$

The first term is a matrix element of the kinetic energy operator of the nuclei, whereas the second term depends on their momenta and is zero when the electronic wavefunction is real. An alternative notation can be introduced:

$$\mathbf{d}^\alpha_{kl}(\mathbf{R}) = \langle\Psi_k(\mathbf{r};\mathbf{R})|\nabla_\alpha|\Psi_l(\mathbf{r};\mathbf{R})\rangle \tag{2.46}$$

$$\mathbf{D}^\alpha_{kl}(\mathbf{R}) = -\langle\Psi_k(\mathbf{r};\mathbf{R})|\nabla^2_\alpha|\Psi_l(\mathbf{r};\mathbf{R})\rangle \tag{2.47}$$

where the terms $\mathbf{d}_{kl}(\mathbf{R})$ and $\mathbf{D}_{kl}(\mathbf{R})$ are called the *first* and *second derivative couplings* respectively. We also refer $\mathbf{d}_{kl}(\mathbf{R})$ as the *nonadiabatic coupling* since, as it can be seen in Eq.(2.45), the diagonal contribution of C_{kk} depends only on a single adiabatic wavefunction $|\Psi_k(\mathbf{r};\mathbf{R})\rangle$:

$$C_{kk} = \langle \Psi_k(\mathbf{r};\mathbf{R})|\hat{T}_N(\mathbf{R})|\Psi_k(\mathbf{r};\mathbf{R})\rangle \tag{2.48}$$

and corresponds to a correction to the (adiabatic) eigen-value E_k of the Schrödinger equation (2.40). The *adiabatic approximation* consists in taking into account only this diagonal elements, reducing the set of Eq.(2.44) to:

$$[\hat{T}_N(\mathbf{R}) + E_k(\mathbf{R}) + C_{kk}(\mathbf{R})]|\chi_k(\mathbf{R},t)\rangle = i\hbar\frac{\partial|\chi_k(\mathbf{R},t)\rangle}{\partial t} \qquad k = 0,1,2,...,\infty \tag{2.49}$$

Neglecting the off-diagonal terms we are assuming that the motion of the nuclei proceeds without changing the quantum state "k" of the electronic subsystem during time evolution. Consequently, the wavefunction Eq.(2.43) is reduced to a single product term:

$$|\Phi(\mathbf{r},\mathbf{R},t)\rangle \approx |\Psi_k(\mathbf{r};\mathbf{R})\rangle|\chi_k(\mathbf{R},t)\rangle \tag{2.50}$$

The final simplification consists in neglecting also the coupling term:

$$[\hat{T}_N(\mathbf{R}) + E_k(\mathbf{R})]|\chi_k(\mathbf{R},t)\rangle = i\hbar\frac{\partial|\chi_k(\mathbf{R},t)\rangle}{\partial t} \tag{2.51}$$

that defines the famous Born-Oppenheimer approximation.

2.6 Born-Oppeheimer molecular dynamics

In this section we will derivate the the semiclassical (or quantum-classical) molecular dynamics from Quantum Mechanics in the particular case of the Born-Oppenheimer approximation. The idea of semiclassical dynamics is to keep the electrons as a quantum objects while treating the atoms as a classical bodies. The formal deduction will be done following the route of Tully [61], that starts by rewriting the corresponding nuclear wavefunction in the following manner [65,78]:

$$|\chi_k(\mathbf{R},t)\rangle = A_k(\mathbf{R},t)\exp(\frac{i}{\hbar}S_k(\mathbf{R},t)) \tag{2.52}$$

which can always be done for any complex function depending on $\mathbf{R}$ and t. $A_k(\mathbf{R},t)$ and $S_k(\mathbf{R},t)$ are the amplitude factor and phase, both considered to be real and $A_k(\mathbf{R},t) > 0$, noticing that $A_k(\mathbf{R},t)^2$ corresponds to the particle probability density $|\chi(\mathbf{R},t)|^2 \equiv \rho_k(\mathbf{R},t)$.

Some algebra is needed in order to introduce Eq.(2.52) into the time-dependent Eq.(2.51) (see appendix A.3 for further details). Once the real and imaginary part are separated, we obtain the following two expressions:

$$\frac{\partial S_k(\mathbf{R},t)}{\partial t} + \sum_{\alpha=1}^{N}\frac{1}{2M_\alpha}(\nabla_\alpha S_k(\mathbf{R},t))^2 + E_k(\mathbf{R}) = \sum_{\alpha=1}^{N}\frac{\hbar^2}{2M_\alpha}\frac{\nabla_\alpha^2 A_k(\mathbf{R},t)}{A_k(\mathbf{R},t)} \tag{2.53}$$

$$\frac{\partial A_k(\mathbf{R},t)}{\partial t} + \sum_{\alpha=1}^{N}\frac{1}{M_\alpha}(\nabla_\alpha A_k(\mathbf{R},t))(\nabla_\alpha S_k(\mathbf{R},t)) + \sum_{\alpha=1}^{N}\frac{1}{2M_\alpha}A_k(\mathbf{R},t)(\nabla_\alpha^2 S_k(\mathbf{R},t)) = 0 \tag{2.54}$$

that are equivalent to the TISE and closely related with Bohmian (fluid) dynamics [79].

Multiplying Eq.(2.54) by $2A_k(\mathbf{R},t)$ we can obtain an expression related with the probability flux, $\mathbf{J}_{k,\alpha}(\mathbf{R},t)$:

$$\frac{\partial \rho_k(\mathbf{R},t)}{\partial t} + \sum_{\alpha=1}^{N} \nabla_\alpha \mathbf{J}_{k,\alpha}(\mathbf{R},t) = 0 \tag{2.55}$$

where:

$$\mathbf{J}_{k,\alpha}(\mathbf{R},t) = \frac{\rho_k(\mathbf{R},t)(\nabla_\alpha^2 S_k(\mathbf{R},t))}{M_\alpha} \tag{2.56}$$

We see that the spatial variation of the phase of the wavefunction characterizes the probability flux; the stronger the phase variation, the more intense the flux.

On the other hand, we can also define a velocity field given by:

$$\mathbf{v}_\alpha = \frac{\mathbf{J}_{k,\alpha}(\mathbf{R},t)}{\mathbf{p}_\alpha} = \frac{S_k(\mathbf{R},t)}{M_\alpha} \tag{2.57}$$

where $\mathbf{p}_\alpha$ is the linear momentum of particle α.

2.6.1 Classical nuclei approximation

Heissemberg's Uncertainty Principle establishes that is not possible to know at the same time two conjugate magnitudes with the same level of accuracy. The precision in one of them runs in favor of the imprecision in the other and viceversa. The error of both measurements is related with the value of $\hbar$. This natural impossibility spoils the idea of classical (deterministic) trajectories and is the underlying origin of the probability character of Quantum Mechanics. However, Classical Mechanics can be derived from Quantum Mechanics imposing the limit $\hbar \to 0$, known as the *classical limit* [80].

In this limit, the right side term of Eq.(2.53) can be neglected since it explicitly depends on $\hbar$. Doing so, we get a similar equation to the Hamiton-Jacobi equation in Classical Mechanics [81]:

$$\frac{\partial S_k(\mathbf{R},t)}{\partial t} + \sum_{\alpha=1}^{N} \frac{1}{2M_\alpha}(\nabla_\alpha S_k(\mathbf{R},t))^2 + E_k(\mathbf{R}) = 0 \tag{2.58}$$

We could consider that, in the classical approximation, $|\chi_k(\mathbf{R},t)\rangle$ describes a fluid of non-interacting classical particles of mass m subjet to the potential $E_k(\mathbf{R})$. We can transform Eq.(2.58) using the equivalences of Eq.(2.57) into:

$$\frac{\partial S_k(\mathbf{R},t)}{\partial t} + \sum_{\alpha=1}^{N} \frac{1}{2} M_\alpha \mathbf{v}_\alpha^2 + E_k(\mathbf{R}) = 0 \tag{2.59}$$

Taking the gradient in both sides:

$$\nabla \frac{\partial S_k(\mathbf{R},t)}{\partial t} + \sum_{\alpha=1}^{N} \frac{1}{2} \nabla_\alpha M_\alpha \mathbf{v}_\alpha^2 = -\nabla E_k(\mathbf{R}) \tag{2.60}$$

Some algebra:

$$\frac{\partial M_\alpha \mathbf{v}_\alpha}{\partial t} + \sum_{\alpha=1}^{N} \mathbf{v}_{k,\alpha} M_\alpha \nabla_\alpha \mathbf{v}_\alpha = -\nabla E_k(\mathbf{R}) \tag{2.61}$$

Lastly, if we assume an incompressible flow:

$$\nabla_\alpha . \mathbf{v}_\alpha = 0 \tag{2.62}$$

the second term on the left of Eq.(2.61) vanish, and then we obtain a set of Newton equations:

$$M_\alpha \ddot{\mathbf{R}}_\alpha = -\nabla_\alpha E_k^{BO}(\mathbf{R}) \qquad \alpha = 1, 2, 3, ..., N \tag{2.63}$$

separately for each decoupled electronic state k.

Reading Eq.(2.63), the nuclei move according to classical mechanics in an effective potential, E_k^{BO}. Thus, the force acting on the nuclei α is obtained by the derivation of the PES that comes from solving the simultaneously the TISE (2.40) of clamped nuclei for the "kth" state at the given nuclear configuration $\{\mathbf{R}(t)\}$. Applying the HF theorem (2.22) we get a final expression of the force acting on each nuclei:

$$M_\alpha \ddot{\mathbf{R}}_\alpha = -\langle \Psi_k(\mathbf{r};\mathbf{R}) | \nabla_\alpha \hat{H}_e(\mathbf{r},\mathbf{R}) | \Psi_k(\mathbf{r};\mathbf{R}) \rangle \qquad \alpha = 1, 2, 3, ..., N \tag{2.64}$$

2.7 Non-adiabatic dynamics in Ehrenfest formalism

The Born-Oppenheimer approximation can be applied for a great number of systems however the non-adiabaticity is presented in many others (charge transfer, photoisomerization reactions, etc.) in which the change in the nuclear configurations might vary the adiabatic states populations. Such a distortion of the electron cloud will, in turn, influence the nuclear trajectory. To introduce the electronic-nuclear feedback the quantum-mechanical time evolution of the electrons must be maintained. The simplest way of doing this is given by the so-called *Mean field method* or *Ehrenfest method*. It starts by using the simplest possible product ansatz that regards the separation between the nuclear and electronic contributions to the total wavefunction [61, 82]:

$$|\Phi(\mathbf{r},\mathbf{R},t)\rangle \approx |\Psi(\mathbf{r},t)\rangle |\chi(\mathbf{R},t)\rangle exp[-\frac{i}{\hbar}\int E_k(\mathbf{R})dt] \tag{2.65}$$

where the electronic and nuclear wavefunctions are separately normalized to unity at every instant of time, and the exponent is a convenient phase factor chosen in such way that the final expression is easy to handle [60].

Substituting the Eq.(2.65) of above into the TDSE (2.1), we obtain two expressions:

$$i\hbar \frac{\partial |\Psi(\mathbf{r},t)\rangle}{\partial t} = -\sum_{i=1}^{n} \frac{\hbar^2}{2m_e} \nabla_i^2 |\Psi(\mathbf{r},t)\rangle + \{\langle \chi(\mathbf{R},t) | \hat{V}_{n-e}(\mathbf{r},\mathbf{R}) | \chi(\mathbf{R},t)\rangle\} |\Psi(\mathbf{r},t)\rangle \tag{2.66}$$

$$i\hbar \frac{\partial |\chi(\mathbf{R},t)\rangle}{\partial t} = -\sum_{\alpha=1}^{N} \frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 |\chi(\mathbf{R},t)\rangle + \{\langle \Psi(\mathbf{r},t) | \hat{H}_e(\mathbf{r},\mathbf{R}) | \Psi(\mathbf{r},t)\rangle\} |\chi(\mathbf{R},t)\rangle \tag{2.67}$$

that are the basics of the Time-dependent self-consistent field (TDSCF) introduced by Dirac [83] in 1930. First applications to atomic an molecular system came, however, much later in the 70-80's [84–90]. Noticing that in this scheme both electrons and nuclei move quantum-mechanically in time-dependent effective potentials (given by the expressions in the braces). The assumption of a single-determinant ansatz like Eq.(2.65) produces an anticipated mean-field description of the coupled nuclear-electronic dynamics.

Using the same equation for the nuclear wavefunction (2.52) (and following almost the same algebra as in appendix A.3), we can obtain one expression in terms of phase from Eq.(2.67):

$$\frac{\partial S_k(\mathbf{R},t)}{\partial t} + \sum_{\alpha=1}^{N} \frac{1}{2M_\alpha}(\nabla_\alpha S_k(\mathbf{R},t))^2 + \langle\Psi(\mathbf{r},t)|\hat{H}_e(\mathbf{r},\mathbf{R})|\Psi(\mathbf{r},t)\rangle = \\ = \sum_{\alpha=1}^{N} \frac{\hbar^2}{2M_\alpha}\frac{\nabla_\alpha^2 A_k(\mathbf{R},t)}{A_k(\mathbf{R},t)} \tag{2.68}$$

The classical limit can be also introduced here as we did in Sec.2.6.1, obtaining a final expression:

$$M_\alpha \ddot{\mathbf{R}}_\alpha(t) = -\nabla_\alpha V^E(\mathbf{R}(t)) = -\nabla_\alpha \langle\Psi(\mathbf{r},\mathbf{R})|\hat{H}_e(\mathbf{r},\mathbf{R})|\Psi(\mathbf{r},\mathbf{R})\rangle \qquad \alpha = 1,2,3,...,N \tag{2.69}$$

where V^E is often called the Ehrenfest potential. Noticing that $\mathbf{R}(t)$ represents the classical-mechanical nuclear positions and no the quantum ones (given by the full nuclear wavefunction $|\chi(\mathbf{R},t)\rangle$). If we use the HF theorem for time-dependent wavefunctions introduce in Sec.2.3.2:

$$M_\alpha \ddot{\mathbf{R}}_\alpha(t) = -\langle\Psi(\mathbf{r},\mathbf{R})|\nabla_\alpha \hat{H}_e(\mathbf{r},\mathbf{R})|\Psi(\mathbf{r},\mathbf{R})\rangle \qquad \alpha = 1,2,3,...,N \tag{2.70}$$

we get an easier expression to implement for the nuclear force acting on the nuclei α.

In the classical limit, first TDSCF Eq.(2.66) can be reduced to:

$$i\hbar\frac{\partial}{\partial t}|\Psi_k(\mathbf{r},\mathbf{R})\rangle = \hat{H}_e|(\mathbf{r},\mathbf{R})|\Psi_k(\mathbf{r},\mathbf{R})\rangle \tag{2.71}$$

if the quantum nuclear positions are replaced by the classical positions (see Ref. [75]).

The Ehrenfest molecular dynamics that relies on solving Newtons equation for the nuclei, Eq.(2.70), simultaneously with Schrödinger equation for the electrons, Eq.(2.71), is often called Ehrenfest MD in honor of Paul Ehrenfest who was the first to address the essential question of how Newtonian classical dynamics of point particles can be derived from Schrödinger time-dependent wave equation [91].

2.7.1 Time-dependent adiabatic couplings, and nuclear forces

The total wavefunction given by Eq.(2.65) can be written as a linear combination of adiabatic eigenfunctions:

$$|\Psi(\mathbf{r},t)\rangle = \sum_{l=0}^{\infty} a_l|\Psi_l(\mathbf{r};\mathbf{R})\rangle exp[-\frac{i}{\hbar}\int E_k(\mathbf{R})dt] \tag{2.72}$$

that are solutions of the TISE. Insertion of this ansatz into the time-dependent electronic Eq. (2.71) followed by multiplication from the left by $\langle\Psi_l(\mathbf{r};\mathbf{R})|$ and integration over the electronic coordinates leads to a set of coupled differential equations:

$$\frac{da_k}{dt} = \sum_{l=0}^{\infty} a_l C_{kl} exp[-\frac{i}{\hbar}\int (E_l(\mathbf{R}) - E_k(\mathbf{R})dt] \tag{2.73}$$

where:

$$C_{kl} = \langle\Psi_l(\mathbf{r};\mathbf{R})|\frac{\partial}{\partial t}|\Psi_l(\mathbf{r};\mathbf{R})\rangle \tag{2.74}$$

are the nonadiabatic coupling elements, that can be related with the nonadiabatic coupling vector by using the chain rule:

$$C_{kl} = \frac{d\mathbf{R}}{dt}\langle\Psi_k(\mathbf{r};\mathbf{R})|\nabla|\Psi_l(\mathbf{r};\mathbf{R})\rangle = \frac{d\mathbf{R}}{dt}\mathbf{d}_{kl} \tag{2.75}$$

Integration of Eq.(2.73) yields the expansion coefficients $a_k(t)$ whose square modulus, $|a_k(t)|^2$, can be interpreted as the probability of finding the system in the adiabatic state k at time t.

Using the definition of the Ehrenfest potential and inserting Eq.(2.72) is quite straightforward to deduce that:

$$V^E(\mathbf{R}(t)) = \langle\Psi(\mathbf{r},\mathbf{R})|\hat{H}_e(\mathbf{r},\mathbf{R})|\Psi(\mathbf{r},\mathbf{R})\rangle = |a_k(t)|^2\sum_{k=0}^{\infty}E_k(\mathbf{R}) \tag{2.76}$$

meaning that the atoms evolve on an effective potential representing an average over the adiabatic states weighted by their state populations $|a_k(t)|^2$.

Following the Ref. [92], nuclear forces of Eq.(2.70) can be also expressed in terms of probabilities:

$$\mathbf{F}_\alpha = -\sum_{k=0}^{\infty}|a_k(t)|^2\nabla_\alpha E_k(\mathbf{R}) + \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}a_k(t)^* a_l(t)(E_l(\mathbf{R}) - E_k(\mathbf{R}))d_{lk}^\alpha \tag{2.77}$$

The Ehrenfest method has been applied with great success to a number of chemical problems including energy transfer at metal surfaces, the study of excited state lifetimes of organic molecules and currents in nanowires [93–95]. The classical description of the heavy ions provides an intuitive way to understand their motion and is readily handled using molecular dynamics. Another advantage is that the electrons can be treated accurately using a single-particle approach, such as TD-DFT theory [33, 96–99]. However, due to its mean-field character the method has some serious limitations. A system that was initially prepared in a pure adiabatic state will be in a mixed state when leaving the region of strong nonadiabatic coupling [92]. For our practical purposes in the study of double ionizations of biomolecules, it means that the initial excited state propagated by TD-DFT Ehrenfest MD will never relax to the ground state density. A switching with a BO surface will be necessary instead.

2.8 Car-Parrinello molecular dynamics

Aiming to cut down the computational expenses of molecular dynamics, Car and Parrinello [100] introduced a non-obvious approach in which the electronic wavefunctions are accounted as dynamical variables. It can be considered to combine the advantages of both Ehrenfest and BornOppenheimer molecular dynamics in an optimal way.

Briefly, in Ehrenfest dynamics the time step to integrate Eq.(2.70) and Eq.(2.71) simultaneously is dictated by the later, i.e., by the intrinsic dynamics of the electrons (much faster than nuclear motion). Contrary to that, there is no electron dynamics involved in solving the BO equations of motion, Eq.(2.40) and Eq.(2.64), because the electronic problem is treated within the TISE. This implies that these equations of motion can be integrated on the time scale given by nuclear motion, thus allows us to use a larger molecular dynamics time step. However, this means that the electronic structure problem has to be solved self-consistently at each molecular dynamics step, whereas this is avoided in Ehrenfest dynamics due to the possibility of propagating the wave function simply by applying the Hamiltonian to an initial wave function (obtained by a single self-consistent optimization at the very beginning of such a simulation).

In this sense, Car-Parrinello (CP) approach integrates the equations of motion on the (long) time scale set by the nuclear motion, and takes advantage of the smooth time evolution of the dynamically evolving electronic subsystem (avoiding the minimization each time step of dynamic). A Lagrangian is defined for the electronic system as follows:

$$L = \sum_i \mu \langle \dot{\psi}_i | \dot{\psi}_i \rangle - E[\{\psi_i\}, \{\mathbf{R}_I\}, \{\alpha_n\}] \tag{2.78}$$

where μ is a fictitious mass associated with the electronic wavefunctions, E is the Kohn-Sham energy functional, $\mathbf{R}_I$ is the position of the ion I, and α_n define the size and shape of the unit cell. The kinetic-energy term in the Lagrangian is due to the fictitious dynamics of the electronic degrees of freedom. The Kohn-Sham energy functional takes the place of the potential energy in a conventional Lagrangian formulation. The electronic wavefunctions are subject to the constrains of orthonoramality:

$$\langle \psi_i(\mathbf{r}) | \psi_j(\mathbf{r}) \rangle = \delta_{ij} \tag{2.79}$$

The corresponding Newtonian equations of motion are obtained form the associated Eurler-Lagrange equations [81] but in terms of nuclear positions and orbitals:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{R}}} = \frac{\partial L}{\partial \mathbf{R}_I} \tag{2.80}$$

$$\frac{d}{dt}\frac{\delta L}{\delta \dot{\psi}_i^*} = \frac{\delta L}{\delta \dot{\psi}_i^*} \tag{2.81}$$

The second equation involves functional derivatives because the orbitals are continuous scalar fields. In practice, however, these fields are represented in a basis (e.g. on a discrete real-space mesh or in terms of plane wave components), and the concept of functional derivation reduces to the partial derivation with respect to the components of the field in the basis. By performing the operations indicated above, we arrive at the Car-Parrinello equations of motion:

$$M_I \ddot{\mathbf{R}}_I(t) = \frac{\partial}{\partial \mathbf{R}_I} \langle \Psi_0(\mathbf{r}) | \hat{H}_e(\mathbf{r}, \mathbf{R}) | \Psi_0(\mathbf{r}) \rangle + \frac{\partial}{\partial \mathbf{R}_I} \{constrains\} \tag{2.82}$$

$$\mu \ddot{\psi}_i(t) = \frac{\delta}{\delta \psi_i^*} \langle \Psi_0(\mathbf{r}) | \hat{H}_e(\mathbf{r}, \mathbf{R}) | \Psi_0(\mathbf{r}) \rangle + \frac{\delta}{\delta \psi_i^*} \{constrains\} \tag{2.83}$$

These equations can be particularized to one particle Hamiltonians such as those resulting from Kohn-Sham theory (described in next section). The proper orbital orthonormality Eq.2.79 must be imposed by Lagrange multipliers Λ_{ij}:

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \langle \Psi_0(\mathbf{r}) | \hat{H}_e^{KS}(\mathbf{r}, \mathbf{R}) | \Psi_0(\mathbf{r}) \rangle \tag{2.84}$$

$$\mu \ddot{\psi}_i(t) = -\hat{H}_e^{KS}(\mathbf{r}, \mathbf{R}) | \psi_i(\mathbf{r}) \rangle + \sum_j \Lambda_{ij} | \psi_j(\mathbf{r}) \rangle \tag{2.85}$$

Chapter 3

Density Functional Theory

Density functional theory (DFT) is not based on the wavefunction but on the electron probability density function (or electron density function) $n(\mathbf{r})$. Given a normalized electronic wavefunction $|\Psi(\mathbf{r})\rangle$, the probability of finding an electron in the space-spin volume element $d\mathbf{r}_1$ located in $\mathbf{r}_1$ while simultaneously another electron is in $d\mathbf{r}_2$ at $\mathbf{r}_2$ and so on, is giving by the product $\Psi^*(\mathbf{r}_1...\mathbf{r}_n)\Psi(\mathbf{r}_1...\mathbf{r}_n)d\mathbf{r}_1...d\mathbf{r}_n$. If we are interested only in probability of finding one electron in $d\mathbf{r}_1$ at $\mathbf{r}_1$, independent of where the other electrons are, then we must average over all space-spin coordinates of the other electrons. Thus the electron density for a single electron $n(\mathbf{r})$ is defined as the following multiple integral[1]:

$$n(\mathbf{r}) = N \int ... \int |\Psi(\mathbf{r}_1, \mathbf{r}_2...\mathbf{x}_N)|^2 d\mathbf{x}_2 d\mathbf{x}_3...d\mathbf{x}_N \tag{3.1}$$

As the electrons are indistinguishable, the integral of $n(\mathbf{r})$ over the volume element $d\mathbf{r}$ determines the probability of finding any of the N electrons within that volume:

$$\int n(\mathbf{r}) d\mathbf{r} = Prob. \tag{3.2}$$

If the integral is extended to the whole space:

$$\int n(\mathbf{r}) d\boldsymbol{\tau} = N \tag{3.3}$$

Unlike the wavefunction, the electronic density is measurable (i.e. by X-ray diffraction or electron diffraction) and also is a non-negative function of only three spatial variables which vanishes at infinity. No matter how big the molecule may be, the electron density remains a function of three variables while the complexity of the wavefunction is increased with the number of electrons (exactly it depends on $6n$ spin and spatial variables). Another advantage is that in about the same time needed for a Hartree-Fock calculation, DFT can often obtain results about the same quality as from MP2 calculations [101]. As we will see later, the flip side of using DFT is the unknown form of the *exchange-correlation functional*, E_{XC} [102]. Approximations are necessary to make up this term, but unlike the wavefunction methods where the solutions can be improved by going to bigger basis sets or higher correlation levels, in DFT there is no way to systematically improve the functional. The choice of the functional really holds the success or the failure of the calculation.

[1] In this chapter the bracket notation will be exchanged for the traditional integral notation.

The average electronic energy can be expressed as a functional of the ground state electron density:

$$E[n] = T[n] + V_{Ne}[n] + V_{ee}[n] \tag{3.4}$$

where it has been separated into three functionals in turn, one associated with the average kinetic electron energy, one with the average nuclear-electron attraction energy and another with the average electron-electron repulsion energy.

For any closed-shell system:

$$\hat{V}_{Ne} = -\sum_{\alpha=1}^{N}\sum_{i=1}^{2n}\frac{Z_\alpha}{r_{\alpha i}} = \sum_{i=1}^{2n}\nu(r_i) \tag{3.5}$$

where $\nu(r_i)$ is the *external potential* created by all the nuclei over the i-electron. It can be shown [103] that its functional density form of the nuclear-electron potential is known and perfectly determinate by:

$$V_{Ne}[n] = \int n(\mathbf{r})\nu(\mathbf{r})d\mathbf{r} \tag{3.6}$$

The Hohenberg-Kohn functional $F_{HK}[n]$ is defined as:

$$F_{HK}[n] = T[n] + V_{ee}[n] \tag{3.7}$$

Thus substituting the equations (3.6) and (3.7) into the equation (3.4) we get:

$$E[n] = \int n(\mathbf{r})\nu(\mathbf{r})d\mathbf{r} + F_{HK}[n] \tag{3.8}$$

3.0.1 Hohenberg-Kohn theorems

The Hohenberg-Kohn (HK) theorems [104, 105] are the base of the density functional theory and can be summarized as follows:

1) For an external potential $\nu(\mathbf{r})$ there is a unique ground-state non-degenerate wavefunction $|\Psi_0(\mathbf{r})\rangle$ that gives rise to a unique ground state density $\mathrm{n}_0(\mathbf{r})$. The first HK theorem states that this mapping can be done backwards too, that is, for any v-representable electronic density[2] $\mathrm{n}(\mathbf{r})$, the external potential which gives rise to it, is unique. This can be expressed by the following bijective map:

$$\begin{array}{c} \nu(\mathbf{r}) \rightarrow |\Psi_0(\mathbf{r})\rangle \rightarrow n_0(\mathbf{r}) \\ \longleftarrow \\ HK \quad theorem \end{array} \tag{3.9}$$

All the observable properties of a molecule in a ground electronic state are then determined by the ground-state electron density function $n_0(\mathbf{r})$. For instance, the energy E_0, we could represent it as:

$$E_0 = E_0[n_0] \tag{3.10}$$

[2] A electron density is v-representable only if it is the ground state of some external potential.

2) The second Hohenberg-Kohn theorem says that any trial electron density function $n(\mathbf{r}_{prob})$ will give an energy higher than or equal to (if it were the exact electron density function) the true ground state energy:

$$E[n_{prob}] \geq E_0[n_0] \tag{3.11}$$

Similar to the variational theorem for wavefunction-based methods (appendix A.4.1), the second Hohenberg-Kohn theorem assures that we can keep choosing different densities and those that provide lower energies will be closer to the correct.

3.0.2 The Kohn-Sham method

Following the first Hohenberg-Kohn theorem we can rewrite Eq.(3.8) as follows:

$$E[n_0] = \int n_0(\mathbf{r})\nu(\mathbf{r})d\mathbf{r} + F_{HK}[n_0] \tag{3.12}$$

showing the relation between the energy and the ground-state electron density. As it is suggested by the second HK theorem, finding the ground-state can be done by minimizing this expression. Unfortunately, the process is not simple because we do not know the relationship between $F_{HK}[n_0]$ and the density. The key point of the Kohn-Sham method is to consider an auxiliary system of N non-interacting electrons for estimating the kinetic energy of the real (interacting) system [105–108].

The Hamiltonian (in atomic units) for a N non-interacting electrons moving under an external potential $\nu_s(\mathbf{r})$ is given by:

$$\hat{H}_s = \sum_{i=1}^{n} \hat{h}_s(\mathbf{i}) = -\frac{1}{2}\sum_{i=1}^{n} \nabla^2(\mathbf{i}) + \sum_{i=1}^{n} \nu_s(\mathbf{i}) \tag{3.13}$$

And the wavefunction is a single Slater determinant (see appendix A.4.2):

$$|\Psi_s(\mathbf{r})\rangle = (N!)^{-1/2}|\psi_1(\mathbf{r})\bar{\psi_1}(\mathbf{r})\psi_2(\mathbf{r})\bar{\psi_2}(\mathbf{r})...\psi_N(\mathbf{r})\bar{\psi_N}(\mathbf{r})\rangle \tag{3.14}$$

Each $|\psi_i(\mathbf{r})\rangle$ represents a Kohn-Sham spin-orbital. If we suppose that the system is a closed-shell, then the density will be:

$$n_s(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2 \tag{3.15}$$

where the subindex i runs only over the occupied orbitals. At the same time, the kinetic energy functional of the system (using compact Dirac notation) is given by:

$$T_s[n_s] = -\sum_{i=1}^{N} \langle\psi_i(\mathbf{r})|\frac{1}{2}\nabla^2|\psi_i(\mathbf{r})\rangle \tag{3.16}$$

We can impose the condition that Kohn-Sham orbitals will be chosen in such a way that its ground state electron density $n_s(\mathbf{r})$ is exactly the same as the real ground state system: $n_s(\mathbf{r}) = n_0(\mathbf{r})$ (hereafter simply n($\mathbf{r}$)). For the real system, we can rewrite the expression of the energy (3.4) in terms of quantities of the reference non-interacting system:

$$E[n] = T_s[n] + V_{Ne}[n] + V_{ee}[n] + \triangle T[n] + \triangle V_{ee}[n] \tag{3.17}$$

The first and the second term represent the kinetic energy and the nuclear-electron interaction for a non-interacting system, given by equations (3.16) and (3.6) respectively. The third term is the classical electrostatic repulsion energy term, also called Hartree term:

$$V_{ee}[n] = \frac{1}{2}\int\int \frac{n(\mathbf{r}_1)n(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 \tag{3.18}$$

If we knew the set of Kohn-Sham spin-orbitals $\{|\psi_i\rangle\}$ that determinate the ground state density $n(\mathbf{r})$, then the value of these three functionals are readily calculated. By contrast, the fourth and fifth terms of the equation (3.17) involve unknown quantities: the former includes all the deviations of the real kinetic functional from the reference system, while the latter includes the corrections to electron classical self-interaction (for instance, in Eq.(3.18) any of the electrons is forced to repel itself).

Both corrections are grouped within the exchange-correlation functional, E_{XC}:

$$E_{XC}[n] = \triangle T[n] + \triangle V_{ee}[n] \tag{3.19}$$

Substituting the equations (3.6), (3.18) and (3.19) into the equation (3.17) leads to the *Kohn-Sham energy equation*:

$$E[n] = T_s[n] + \int n(\mathbf{r})\nu(\mathbf{r})d\mathbf{r} + \frac{1}{2}\int\int \frac{n(\mathbf{r}_1)n(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 + E_{XC}[n] \tag{3.20}$$

that is exact if we knew the density function and the exchange-correlation energy functional $E_{XC}[n]$.

3.0.3 The exchange-correlation functional

As it has been previously mentioned, the main problem of DFT is the unknown exchange-correlation energy functional $E_{XC}[n]$. The dependence of this functional on the electron density is normally expressed as an interaction between the electron density and "*energy density*" functional $\epsilon_{XC}[n]$:

$$E_{XC}[n] = \int n(\mathbf{r})\epsilon_{XC}[n]d\mathbf{r} \tag{3.21}$$

Taking the expression (3.21) and deriving it with respect electron density, we obtain the exchange correlation potential $\nu_{XC}(\mathbf{r})$,:

$$\nu_{XC}(\mathbf{r}) = \frac{\partial E_{XC}[n]}{\partial n} = \epsilon_{XC}[n] + n(\mathbf{r})\frac{\partial \epsilon_{XC}[n]}{\partial n} \tag{3.22}$$

Many approximations for $E_{XC}[n(\mathbf{r})]$ have been proposed. In the following sections, we will revise a few of them focussing the attention on BLYP functional, which is the one used in this thesis.

Local density approximation (LDA)

The local density approximation [109] represents the simplest approximation to $E_{XC}[n]$. It derives from analysis of the uniform electron gas assuming that the local exchange functional can be derived from the exact solution of this model. In this approximation the energy density

functional $\epsilon_{XC}[n(\mathbf{r})]$ is always treated as a sum of individual exchange and correlation contributions:

$$\epsilon_{XC}[n] = \epsilon_X[n] + \epsilon_C[n] \tag{3.23}$$

The first term of equation (3.23) is:

$$\epsilon_X[n] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}}(n(\mathbf{r}))^{\frac{1}{3}} \tag{3.24}$$

If the previous expression is employed without any correlation functional is known as "X_α *method*" and was proposed by Slater [110]. Vosko, Wilk and Nusair (VWN) obtained an expression for the correlation energy $\epsilon_C[n(\mathbf{r})]$ of a uniform electron gas using Monte Carlo calculations [111].

$$\epsilon_C[n] = \epsilon_C^{VWN}[n] \tag{3.25}$$

Finally, using the equations (3.24) and (3.25), the expression of exchange-correlation functional within LDA approach can be written as:

$$E_{XC}^{LDA}[n] = \int n(\mathbf{r})\epsilon_{XC}[n]d\mathbf{r} = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}}\int (n(\mathbf{r}))^{\frac{4}{3}}d\mathbf{r} \tag{3.26}$$

Already simple approximations such as the local density approximation (LDA) have turned out to be quite successful in the calculation of structures and total energies of molecules and solids [112, 113].

LSD approximation

Local spin density approximation is suited to treat the open shell systems or molecules next to dissociation. While in the LDA two paired electrons with opposite spin have the same spatial Kohn-Sham orbital $|\psi(\mathbf{r})\rangle$, in the LSDA they occupied different spatial orbitals $|\psi_\alpha(\mathbf{r})\rangle$ and $|\psi_\beta(\mathbf{r})\rangle$. Thus α and β densities are separately minimized:

$$n_r(\mathbf{r}) = n_r^\alpha(\mathbf{r}) + n_r^\beta(\mathbf{r}) = \sum_{i=1}^{N_\alpha}|\psi_{\alpha i}(\mathbf{r})|^2 + \sum_{i=1}^{N_\beta}|\psi_{\beta i}(\mathbf{r})|^2 \tag{3.27}$$

Thus we define two different exchange-correlation functionals:

$$E_{XC}^{LSDA}[n^\alpha] = E_X^{LSDA}[n^\alpha] + E_C^{LSDA}[n^\alpha, n^\beta] \tag{3.28}$$

$$E_{XC}^{LSDA}[n^\beta] = E_X^{LSDA}[n^\beta] + E_{C'}^{LSDA}[n^\alpha, n^\beta] \tag{3.29}$$

Like in LDA the exchange-correlation functional $E_{XC}^{LDA}[n]$ and its derivate $\nu_{XC}^{LDA}(\mathbf{r})$, $E_{XC}^{LSDA}[n^\alpha]$ and $E_{XC}^{LSDA}[n^\beta]$ and their functional derivatives can be accurately calculated [107, 114].

Generalized gradient approximation (GGA)

In the local density approximation, the functional used to obtain the exchange correlation energy only depends on the electron density. An improvement of this method is to consider exchange

correlation functionals that not only involve the electron density $n(\mathbf{r})$ but also its local gradients (first derivate with respect to position).

The introduction of gradient correction terms to the LDA functional by Perdew and Yang [115] lead to the *Generalized Gradient Approximation* (GGA) and the *gradient-corrected* functionals. These functional are also known as *nonlocal* functionals, in contrast to the local LDA and LSDA functionals, since the gradient is the change over an infinitesimal distance beyond the "local" point of the coordinates $\mathbf{r}$. Among all the GGA functionals it is worth mentioning the B88 [116], G96 [117] and P86 [118] functionals. Lee, Yang and Parr developed an other well-known correlation functional [119] (LYP) that is a density gradient expansion, based on the orbital functional for the correlation energy of Colle and Salvetti [120]. The combination of such functional with gradient corrections to the LDA method - specifically the exchange correction of Becke (B88) forms the so-called BLYP functional used in the MD simulation of the present work.

The GGA functionals represent a noticeable improvement upon LDA especially for molecular systems but also work very well for atoms and solids [121]. They can in particular reproduce well hydrogen bonded system like liquid water [122]. Other types of functionals are the hybrid functionals [123] that incorporate a fraction of Hartree-Fock exchange.

3.0.4 Solutions of the Kohn-Sham equations

After discussing the foundations of DFT, we will shortly describe methods for solving the Kohn-Sham problem. There are basically two methods to find the ground-state density within the Kohn-Sham framework for a given external potential [56]. The first one focuses on the self-consistent resolution of the equation obtained from the derivation of energy functional (3.20) to respect a small change in the electron density. The second one aims at directly minimizing Eq.(3.20) through the imposition of some constraints.

Diagonalization technique

From the second HK theorem we can extract a condition for an electronic density to minimize the energy functional:

$$\delta E_v[n] = \delta T_s[n] + \int d\mathbf{r} \left(\nu(\mathbf{r}) + \frac{\partial V_{ee}[n]}{\partial n(\mathbf{r})} + \nu_{XC}(\mathbf{r}) \right) \delta n(\mathbf{r}) = 0 \tag{3.30}$$

Where:

$$\frac{\partial V_{ee}[n]}{\partial n(\mathbf{r})} = \int \frac{n(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 = \nu_J(\mathbf{r}) \tag{3.31}$$

is the Hartree potential.

Attending now to the first HK theorem, for a given density, n($\mathbf{r}$), we can associate the so-called Kohn-Sham potential, $\nu_{KS}(\mathbf{r})$, that makes the function to be the ground-state density of the non-interacting electron system. The variation of the kinetic functional, $\delta T_s[n]$, can be expressed in terms of this potential:

$$\delta T_s[n] = -\int d\mathbf{r} \nu_{KS}(\mathbf{r}) \delta n(\mathbf{r}) \tag{3.32}$$

Introducing Eq.(3.32) into Eq.(3.30) we get:

$$\delta E_v[n] = -\int d\mathbf{r} \nu_{KS}(\mathbf{r}) \delta n(\mathbf{r}) + \int d\mathbf{r} \left(\nu(\mathbf{r}) + \int \frac{n(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 + \nu_{XC}(\mathbf{r}) \right) \delta n(\mathbf{r}) = 0 \tag{3.33}$$

and the condition for $\mathrm{n}(\mathbf{r})$ to be the minimum of $E_v[n]$ is:

$$\nu_{KS}(\mathbf{r}) = \nu(\mathbf{r}) + \nu_J(\mathbf{r}) + \nu_{XC}(\mathbf{r}) \tag{3.34}$$

Eq.(3.34) means that finding an electronic density $\mathrm{n}(\mathbf{r})$ to be the ground-state of an external potential $\nu(\mathbf{r})$, it is equivalent to find the ground state density of the non-interacting electron problem in the external potential constructed by the right hand side expression of Eq.(3.34). In fact, the two potentials are equal if and only if they lead to the same ground state electronic density (HK theorem).

The diagonalization technique starts from a guess density $\mathrm{n}(\mathbf{r}) = \mathrm{n}_{old}(\mathbf{r})$ and constructs the associated Kohn-Sham potential:

$$\nu_{KS}^{old}(\mathbf{r}) = \nu(\mathbf{r}) + \int \frac{n^{old}(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 + \nu_{XC}(\mathbf{r}; n^{old}) \tag{3.35}$$

Then the Schrödinger equation for the non-interacting electrons:

$$-\frac{1}{2}\nabla_i^2 |\psi_i^{new}(\mathbf{r})\rangle + \nu_{KS}^{old}(\mathbf{r})|\psi_i^{new}(\mathbf{r})\rangle = \epsilon_i^{new}|\psi_i^{new}(\mathbf{r})\rangle \tag{3.36}$$

is solved by the orbital expansion in a basis set:

$$|\psi_i(\mathbf{r})\rangle = \sum_{n=1}^{m} c_{ni}|\phi_n(\mathbf{r})\rangle \tag{3.37}$$

The resultant set of equations can be solved by diagonalization. The process can be performed by explicit construction of the matrix $(\mathrm{H}_{\alpha\beta})$ which is then diagonalized by standard methods when the basis set is not too large. For the case of large systems and/or large basis sets, other iterative techniques are used. The Lanczos method [124–127], for example, avoids the explicit construction of the Kohn-Sham matrix: it is sufficient in these methods to have a procedure to apply (successively) the Kohn-Sham matrix on vectors $\{c_\alpha\}$. Only vectors $\{c_\alpha\}$ need then to be stored.

Once we have solved the non-interacting electron problem in the external potential $\nu_{KS}^{old}(\mathbf{r})$, to obtain one electron orbitals $\{|\psi_i\rangle\}$ (expanded on the basis set) we construct the ground state density of the non-interacting electron system simply as:

$$n^{new}(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i^{new}(\mathbf{r})|^2 \tag{3.38}$$

After that is checked if $\mathrm{n}_{new} = \mathrm{n}_{old}$, if this is not the case, the procedure is iterated with a new guess density constructed from n_{new}.

Constrained minimization technique

Now we are looking for an electronic density $n(\mathbf{r})$ that minimizes the functional for the Kohn-Sham energy (3.20), but assuming the constraint of the spin-orbitals being orthormal:

$$\int \psi_i^*(\mathbf{r})\psi_j(\mathbf{r})d\mathbf{r} = \delta_{ij} \tag{3.39}$$

This problem is solved by the Lagrange's method:

$$L[n_0] = E[n_0] - \mu \left[\sum_{i=1}^{N}\sum_{j=1}^{N} l_{ij} \int \psi_i^*(\mathbf{r})\psi_j(\mathbf{r})d\mathbf{r} \right] \tag{3.40}$$

So that, the problem of finding a minimum energy has change to solve the equation:

$$\delta L[n_0] = 0 \tag{3.41}$$

Deriving the equation (3.40) and diagonalizing the expressions obtained (for detailed description see [107]) one can lead to a set of integro-differential equations known as *Kohn-Sham equations*:

$$\begin{aligned} \hat{f}_{KS}(\mathbf{1})|\psi_1(\mathbf{1})\rangle &= \epsilon_1|\psi_1(\mathbf{1})\rangle \\ \hat{f}_{KS}(\mathbf{1})|\psi_2(\mathbf{1})\rangle &= \epsilon_2|\psi_2(\mathbf{1})\rangle \\ \vdots \quad \vdots \quad \vdots & \\ \hat{f}_{KS}(\mathbf{1})|\psi_n(\mathbf{1})\rangle &= \epsilon_n|\psi_n(\mathbf{1})\rangle \end{aligned} \tag{3.42}$$

where the Kohn-Sham operator $\hat{f}_{KS}$ is:

$$\hat{f}_{KS}(\mathbf{i}) = -\frac{1}{2}\nabla^2(\mathbf{i}) + \nu_{KS}(\mathbf{r}) \tag{3.43}$$

and $\nu_{KS}(\mathbf{r})$ represents the Kohn-Sham potential previously defined as:

$$\nu_{KS}(\mathbf{r}) = \nu(\mathbf{r}) + \int \frac{n(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 + \nu_{XC}(\mathbf{r})$$

The resolution of these Kohn-Sham eigenvalue equations is performed by standard iterative methods like the conjugate gradient method [55, 128] or Direct Inversion in Iterative Space method [129–131] (DIIS). In both the Kohn-Sham orbitals are expanded in terms of basis functions following Eq.(3.37). Thus, the set of equations (3.42) can be rewritten as following matrix equation:

$$\mathbf{F}^{KS}\mathbf{C} = \mathbf{SC}\boldsymbol{\epsilon} \tag{3.44}$$

Being the elements of the Kohn-Sham matrix and overlap matrix:

$$F_{ij}^{KS} = \int \phi_n^*(\mathbf{1})\hat{f}_{KS}(\mathbf{1})\phi_n(\mathbf{1})d\mathbf{r} \tag{3.45}$$

$$S_{ij} = \int \phi_n^*(\mathbf{1})\phi_n(\mathbf{1})d\mathbf{r} \tag{3.46}$$

A operational procedure can be proposed as follows:

1. Choose a basis set, a initial molecular geometry and a exchange-correlation functional.
2. Obtain a guess of the density $n(\mathbf{r})$ and construct an explicit expression for the Kohn-Sham operator $\hat{f}_{KS}$.
3. Calculate the overlap, one-electron and two-electron integrals to obtain the matrix elements F_{ij}^{KS} and S_{ij}.
4. The Kohn-Sham matrix is orthogonalized and diagonalized giving initial guess of the c's in the basis set expansion of Eq.(3.37).
5. Use the new set of c's to obtain a new density function using the Eq.(3.15).

6. Calculate improved matrix elements F_{ij}^{KS} using the new density function and get improved c's and improved density function.

7. Continue the iterative process until the electron density (or another criteria) converge.

It is important to notice that in the DFT approximation only the electronic density, the energy of the Kohn-Sham HOMO orbital [132] and the *chemical potential*, μ:

$$\mu = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + \nu_{eff}(\mathbf{r}) \tag{3.47}$$

have physical sense. On the contrary, the rest of orbital energies and the Kohn-Sham wavefunction do not have a strict physical meaning. In fact, the Kohn-Sham wavefunction, $|\Psi(\mathbf{r})\rangle$, is nothing but a wavefunction that belongs to the set of functions $\{|\psi_i\rangle\}$ whose integral is $n(\mathbf{r})$.

3.0.5 Plane waves basis sets

In this section we will quickly discuss the basis expansion of Eq.(3.37) in the special case of plane wave basis set. Bloch's theorem [133] states that in a periodic solid each electronic wavefunction can be written as the product of a cell-periodic part and a wavelike part:

$$|\psi_i(\mathbf{r})\rangle = \exp(i\mathbf{k}\cdot\mathbf{r})|f_i(\mathbf{r})\rangle \tag{3.48}$$

The cell-periodic part of the wave function can be expanded using a basis set made by a discrete set of plane waves whose wave vectors are reciprocal lattice vectors of the crystal:

$$|f_i(\mathbf{r})\rangle = \sum_{\mathbf{G}} c_{i,\mathbf{G}} \exp(i\mathbf{G}\cdot\mathbf{r}) \tag{3.49}$$

The reciprocal lattice vectors $\mathbf{G}$ are defined by $\mathbf{G}\cdot\mathbf{l} = 2\pi m$ where $\mathbf{l}$ is a lattice vector of the crystal and m is an integer. For example, if we consider an orthorhombic box with lengths L_x, L_y and L_z, the wavevectors $\mathbf{G}$ are:

$$\mathbf{G} = i\frac{2\pi}{L_x}\mathbf{x} + j\frac{2\pi}{L_y}\mathbf{y} + k\frac{2\pi}{L_z}\mathbf{z} \qquad \text{with i, j, k} \in \text{Z} \tag{3.50}$$

A final expression of the electronic wavefunction as a sum of plane waves is obtained if we insert Eq.(3.49) into Eq.(3.48):

$$|\psi_i(\mathbf{r})\rangle = \sum_{\mathbf{G}} c_{i,\mathbf{k}+\mathbf{G}} \exp(i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}) \tag{3.51}$$

Instead of having to solve an infinite number of wavefunctions over all of (infinite) space, the Bloch theorem changes the problem to one of solving a finite number of wavefunctions at an infinite number of possible values for $\mathbf{k}$. In order to simplify the problem to manageable proportions, it is necessary to impose some boundary conditions on the wavefunction, which restrict the allowed values of $\mathbf{k}$. Thus, the system is contained within a supercell which is then replicated periodically throughout space. The supercell must be large enough so that the systems contained within each one, which in reality are isolated, do not interact significantly.

In principle, even if the boundary conditions are applied, an infinite plane wave basis set is required to expand the electronic wavefunctions. In a practical way, since the basis set elements are eigenvectors of the kinetic energy operator:

$$-\frac{1}{2}\nabla_i^2|f_i(\mathbf{r})\rangle = \frac{1}{2}||\mathbf{G}||^2|f_i(\mathbf{r})\rangle \tag{3.52}$$

the size of the basis is determined by inserting a cut-off:

$$E_{kin} = \frac{1}{2}||\mathbf{G}||^2 < E_{cut} \tag{3.53}$$

that includes all plane waves whose kinetic energy is below [55].

Kohn-Sham equations (3.42) can be written in terms of plane-waves using Eq.(3.51):

$$\sum_{\mathbf{G'}} \left[|\mathbf{k}+\mathbf{G}|^2\delta_{\mathbf{GG'}} + \nu(\mathbf{G\text{-}G'}) + \nu_J(\mathbf{G\text{-}G'}) + \nu_{XC}(\mathbf{G\text{-}G'})\right] c_{i,\mathbf{k+G'}} = \epsilon_i c_{i,\mathbf{k+G}} \tag{3.54}$$

where the kinetic energy is diagonal and the various potentials are described in terms of their Fourier transforms [55].

For MD, the main advantage of a plane wave basis set is the independence of the basis set elements with respect to the ionic positions. As it was described in Sec.(2.3), the Hellmann-Feynman theorem can be applied without additional Pulay terms arising from a basis set that would be dependent on the nuclei positions. Therefore, the forces on the ions will be calculated at no extra-cost. There is also no Basis Set Superposition Error for the same reasons. Another advantage of plane wave basis sets is that their quality depends only on the cut-off considered; it is thus easier both to compare results and to make convergence studies with only one number defining the quality of the basis set. Finally, concerning the CPU time, plane wave basis sets have the advantage of being orthonormal and they can not become over-complete (one element being exactly or even approximately a linear combination of other elements, fact that leads to divergences).

However, plane wave basis sets also have disadvantages. The first one is probably the very large number of basis set elements required, specially to cover more localized inner electrons. Indeed calculations might become intratable and it is absolutely necessary to employ pseudopotentials: only valence electrons are considered, not core electrons; in consequence, the electron-ion interactions are not simply the fundamental coulomb attraction but there are also pseudo-potential interactions. Another interesting point is that, not being atom centered orbitals, plane wave basis sets do not easily lead to chemical insight on the electronic structure of the system studied: it is hard to describe the result of a plane wave calculation in a Linear Combination of Atomic Orbital framework.

Other characteristics can be both, an advantage or disadvantage, depending on the application considered. For instance, plane wave basis set elements fill equally the whole simulation box. Thus, regions of high electronic densities near the atomic cores are described at the same level as regions of lower density like the valence region which leads as already mentioned to the necessary use of pseudo-potentials. When an isolated molecule is considered, this turns out to be a further disadvantage as the vacuum region around the molecule is still described at the same level of accuracy. On the contrary, it plays in favor of condensed systems since no region of very low electronic density is present. Another particular situation that makes this feature an advantage is when an electron is localized in a cavity instead of around an atom for example. This is the case for a solvated electron [134] or a solvated silver atom in aqueous solution [135, 136].

3.0.6 Pseudopotential Approximation

As it has been referred, the plane wave basis sets represent a appropriate way to expand electronic wavefunctions on periodic external potentials like crystals or surfaces. However, the very large number of them needed to expand the tightly bound core orbitals and to represent the rapid

oscillations of the valence electron wavefunction in the core rigion[3], make this choice useless when an all-electron calculation is considered.

The Pseudopotential Approximation [137–139] solves both problems by replacing the bare coulomb attraction between electrons and nuclei by a pseudopotential that acts on a set of pseudo wavefunctions (rather than the true valence wavefunctions) within the core region. The most common pseudopotential approach consists of not allowing the relaxation of the core states according to the environment (*frozen core approximation*), although some polarizable core approaches have been proposed [140].

The pseudo-functions proposed are combination of the true and valence wavefunctions:

$$|\widetilde{\psi}_v(\mathbf{r})\rangle = |\psi_v(\mathbf{r})\rangle + \sum_c \alpha_{cv}|\psi_c(\mathbf{r})\rangle \tag{3.55}$$

and satisfy the modified Schrödinger equation:

$$\left[\hat{H} + \sum_c (\epsilon_v - \epsilon_c)|\psi_c(\mathbf{r})\rangle\langle\psi_c(\mathbf{r})|\right] |\widetilde{\psi}_v(\mathbf{r})\rangle = \epsilon_v|\widetilde{\psi}_v(\mathbf{r})\rangle \tag{3.56}$$

where the potential nuclear-electron attraction term of the Hamiltonian operator is replaced by $\hat{V}_{eN} = (Z_c/r)\hat{I}$, being $\hat{I}$ the identity operator. It is possible to construct a pseudo-Hamiltonian:

$$\hat{H}_{PS} = \hat{H} + \sum_c (\epsilon_v - \epsilon_c)|\psi_c(\mathbf{r})\rangle\langle\psi_c(\mathbf{r})| \tag{3.57}$$

with the same eigen values of the original Hamiltonian but smoother, nodeless wavefunction. The associated potential:

$$\hat{V}_{PS} = \frac{Z_c}{r}\hat{I} + \sum_c (\epsilon_v - \epsilon_c)|\psi_c(\mathbf{r})\rangle\langle\psi_c(\mathbf{r})| \tag{3.58}$$

is called a pseudopotential.

It is required however that these functions be transferable, i.e., the pseudopotential for one element can be valid in different calculations for many different chemical environments of this element [141]. A major step forward in the theory of pseudopotentials was the introduction of a "norm-conservation" condition with the proposition by Hamann, Schlüter and Chiang [142] of a set of conditions to ensure transferability of pseudopotentials. Different recipes for constructing pseudopotentials satisfying these conditions have been proposed. They can be either analytic [142, 143] or numeric [144] and also differ by the way valence wave functions are made smooth in the core region.

For these norm-conserving pseudopotentials, a different potential needs to be applied on each orbital depending on its angular component l. These pseudopotentials then have a semi-local form:

$$\hat{V}_{PS}(\mathbf{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \nu_{PS}^{l}(\mathbf{r})|lm\rangle\langle lm| \tag{3.59}$$

where $\nu_{PS}^{l}(\mathbf{r})$ is the pseudopotential corresponding to the angular component l.

[3]These oscillations are a consequence of maintaining the orthogonality of core wavefunctions with the valence wavefunctions.

3.0.7 Localized Molecular Orbitals

For computational purposes, it is convenient to work with Kohn-Sham molecular orbitals, i.e. those that make the Khon-Sham matrix, F^{KS}, diagonal; and that are eigenfunctions of the Kohn-Sham, $\hat{f}_{KS}$, operator at convergence. This corresponds to a specific choice of a unitary (orthogonal) transformation of the occupied molecular orbitals. Once the self-consistent procedure has converged, however, other sets of orbitals may be chosen by forming linear combinations of the canonical molecular orbitals. The total wavefunction, and thus all observable properties, is independent of such a rotation of the molecular orbitals:

$$|\psi'(\mathbf{r})\rangle = U|\psi(\mathbf{r})\rangle \quad \text{where} \quad |\psi'_i(\mathbf{r})\rangle = \sum_{j=1}^{N} u_{ij}|\psi_j(\mathbf{r})\rangle \tag{3.60}$$

The traditional view of molecular bonds is that they are due to an increased probability of finding electrons between two nuclei. The Kohn-Sham orbitals are delocalized over the whole molecule and do not readily reflect this, since the density between two nuclei is the result of many small contributions from many (all) the molecular orbital. The canonical molecular orbitals therefore do not readily allow identification of the bonding properties of the system.

The goal of Localized Molecular Orbitals is to define molecular orbitals that are spatially confined to a relatively small volume, and therefore clearly display which atoms are bonded and furthermore have the property of being approximately constant between structurally similar units in different molecules. A set of localized orbitals may be defined by optimizing the expectation value of a two-electron operator Ω [145, 146]:

$$\langle\Omega\rangle = \sum_{j=1}^{N}\langle\psi'_i(\mathbf{1})\psi'_i(\mathbf{2})|\Omega|\psi'_i(\mathbf{1})\psi'_i(\mathbf{2})\rangle \tag{3.61}$$

In Eq.(3.61), the expectation value of Ω depends on the u_{ij} parameters in Eq.(3.60), that represents again a function optimization problem. The unitary transformation of the orbitals preserves the orthogonality, i.e. the resulting localized molecular orbitals are also orthogonal. Since all observable properties depend only on the total electron density, and not the individual molecular, there is no unique choice for Ω. For example Boys-Foster localization scheme [147] uses the square of the distance between two electrons as the operator, and minimizes the expectation value:

$$\langle\Omega\rangle_{Boys} = \sum_{j=1}^{N}\langle\psi'_i(\mathbf{1})\psi'_i(\mathbf{2})|(r_1 - r_2)^2|\psi'_i(\mathbf{1})\psi'_i(\mathbf{2})\rangle \tag{3.62}$$

This corresponds to determining a set of localized molecular orbitals that minimizes the spatial extent, that is, they are as compact as possible. For systems described by plane wave basis functions the equivalent of the Boys localized orbitals is called Wannier orbitals [148].

3.1 Outline of molecular dynamics simulation based on DFT

In the previous sections we have seen that finding a ground-state density is equal to solve the self-consistent resolution of the electronic Kohn-Sham equations:

$$\left[-\frac{1}{2}\nabla_i^2 + \nu(\mathbf{r}) + \nu_J(\mathbf{r}) + \nu_{XC}(\mathbf{r})\right]|\psi_i(\mathbf{r})\rangle = \epsilon_i|\psi_i(\mathbf{r})\rangle$$

where:

$$\nu(\mathbf{r}) = -\sum_{\alpha=1}^{N} \frac{Z_\alpha}{r_{\alpha i}} \quad ; \quad \nu_J(\mathbf{r}) = \int \frac{n(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 \quad ; \quad \nu_{XC}(\mathbf{r}) = \frac{\partial E_{XC}[n]}{\partial n(\mathbf{r})}$$

are the external, Hartree and exchange-correlation potential respectively.

In a MD simulation this process is combined with the nuclear motion following the general outline:

1. A set of atoms and initial positions and velocities $\{\mathbf{r}_0\}$, $\{\mathbf{v}_0\}$ are read from the input. This file also contains the basis cut-off, time-step, and the proposed exchange-correlation potential, $E_{XC}[n]$.
2. Given Z numbers and positions of the atoms, external potential $\nu(\mathbf{r})$ is set up.
3. A trial electron density n($\mathbf{r}$) is constructed.
4. $\nu_J(\mathbf{r})$ and $\nu_{XC}(\mathbf{r})$ potentials are calculated.
5. Kohn-Sham equations are solved by diagonalization. This is highest CPU time consuming process.
6. With the set of the new Kohn-Sham orbitals, a new election density is constructed.
7. The solution is checked to be consistent, if it is not, go back to step 4.
8. Once a ground electronic density is achieved, the following step is to calculate the atomic forces, $\{\mathbf{f}_0\}$, by the Helmann-Feynman theorem (2.22).
9. From $\{\mathbf{r}_0\}$, $\{\mathbf{v}_0\}$, $\{\mathbf{f}_0\}$, a new set of positions and velocities $\{\mathbf{r}\}$,$\{\mathbf{v}\}$ are obtained by using the Verlet algorithm (detailed in next section) with the time-step read from the input.
10. Positions and velocities are updated and step 2 reinitiated.

3.1.1 Verlet algorithm

The forces are used to calculate the velocities and positions in the next time time step. In molecular dynamics, the most commonly used time integration algorithm is probably the so-called Verlet algorithm [149, 150]. The basic idea is to write two third-order Taylor expansions for the positions $\mathbf{r}(t)$, one forward and one backward in time. Calling $\mathbf{v}$ the velocities, $\mathbf{a}$ the accelerations, and $\mathbf{b}$ the third derivatives of $\mathbf{r}$ with respect to t, one has:

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + ...(t)\Delta t^2 + (1/6)\mathbf{b}(t)\Delta t^3 + O(\Delta t^4) \tag{3.63}$$

$$\mathbf{r}(t-\Delta t) = \mathbf{r}(t) - \mathbf{v}(t)\Delta t + ...(t)\Delta t^2 - (1/6)\mathbf{b}(t)\Delta t^3 + O(\Delta t^4) \tag{3.64}$$

Adding both expressions gives:

$$\mathbf{r}(t+\Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t-\Delta t) + \mathbf{a}(t)\Delta t^2 + O(\Delta t^4) \tag{3.65}$$

This is the basic form of the Verlet algorithm. Since we are integrating Newton's equations, $\mathbf{a}(t)$ is just the force divided by the mass, and the force is in turn a function of the positions $\mathbf{r}(t)$:

$$\mathbf{a}(t) = -(1/m)\nabla V\left(\mathbf{r}(\mathbf{t})\right) \tag{3.66}$$

As one can immediately see, the truncation error of the algorithm when evolving the system by Δt is of the order of Δt^4, even if third derivatives do not appear explicitly. This algorithm is at the same time simple to implement, accurate and stable, explaining its large popularity among molecular dynamics simulators.

A problem with this version of the Verlet algorithm is that velocities are not directly generated. While they are not needed for the time evolution, their knowledge is sometimes necessary. Moreover, they are required to compute the kinetic energy K, whose evaluation is necessary to test the conservation of the total energy E=K+V. This is one of the most important tests to verify that a MD simulation is proceeding correctly. One could compute the velocities from the positions by using

$$\mathbf{v}(t) = \frac{\mathbf{r}(t+\Delta t) - \mathbf{r}(t-\Delta t)}{2\Delta t}. \tag{3.67}$$

However, the error associated to this expression is of order Δt^2 rather than Δt^4. To overcome this difficulty, some variants of the Verlet algorithm have been developed. They give rise to exactly the same trajectory, and differ in what variables are stored in memory and at what times. The leap-frog algorithm, not reported here, is one of such variants [151] where velocities are handled somewhat better. An even better implementation of the same basic algorithm is the so-called velocity-Verlet scheme (used in the CPMD code [152, 153]), where positions, velocities and accelerations at time $t+\Delta t$ are obtained from the same quantities at time t in the following way:

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + (1/2)\mathbf{a}(t)\Delta t^2 \tag{3.68}$$
$$\mathbf{v}(t+\Delta t/2) = \mathbf{v}(t) + (1/2)\mathbf{a}(t)\Delta t \tag{3.69}$$
$$\mathbf{a}(t+\Delta t) = -(1/m)\nabla V\left(\mathbf{r}(t+\Delta t)\right) \tag{3.70}$$
$$\mathbf{v}(t+\Delta t) = \mathbf{v}(t+\Delta t/2) + (1/2)\mathbf{a}(t+\Delta t)\Delta t \tag{3.71}$$

Note how we need 9N memory locations to save the 3N positions, velocities and accelerations, but we never need to have simultaneously stored the values at two different times for any one of these quantities.

3.2 Basics on Time Dependent Density Functional Theory

Up to now we have seen that ground-state DFT is a promising alternative to wavefunction calculations for electronic structure and properties. DFT is in principle exact, built on the Hohenberg-Kohn theorems: the ground-state density of an interacting N-electron system in a static external potential completely determines all of its properties (e.g., energy, bond lengths, even excited states) [105, 107]. The practical power of DFT lies in mapping the system to a fictitious non-interacting system of the same ground-state density, the Kohn-Sham (KS) system, with a one-body potential. Solution of these self-consistent orbital equations is faster than solution of the Schrödinger equation, allowing calculations on large systems of interest in materials science and chemistry.

Time Dependent Density Functional Theory (TD-DFT) is laid in the Runge-Gross theorem [154] that represents an extension of HK theorems for interacting electrons in time-dependent external potentials. A set of equivalent Time-Dependent Kohn-Sham equations (TD-KS) for fictitious noninteracting electrons are also derived. The theory aims to treat for example, atoms and molecules in intense laser fields, electronic transition frequencies, oscillator strengths, polarizabilities, etc. Following Ref. [155], it is convenient to distinguish two regimes in TD-DFT:

(a) If the external time-dependent potential is "small", the complete numerical solution of the time-dependent Kohn-Sham equations can be avoided by the use of linear response theory. This is the case, e.g., for the calculation of photoabsorption spectra. (b) For a "strong" external potential, a full solution of the time-dependent Kohn-Sham equations is in needed. This situation is encountered, for instance, when matter interacts with intense laser fields.

In this survey we will introduce some the basics concepts on TD-DFT focusing our attention on the time propagation of the TD-KS equations in the Ehrenfest MD used in this thesis [33]. Further reading and detailed applications of TD-DFT and Linear Response Theory, can be found in Refs. [156–162].

3.2.1 Runge-Gross theorem

Considering N non-relativistic electrons, interacting via Coulomb repulsion, in a time-dependent external potential, the Runge-Gross theorem [154] states that the densities n(**r**,t) and n'(**r**,t) evolving from a common initial state $|\Psi_0\rangle = |\Psi(t=0)\rangle$ under the influence of two external potentials $\nu(\mathbf{r},t)$ and $\nu'(\mathbf{r},t)$ (both Taylor expandable about the initial time 0) are always different provided that the potentials differ by more than a purely time-dependent (r-independent) function:

$$\nu(\mathbf{r},t) - \nu'(\mathbf{r},t) \neq c(t) \tag{3.72}$$

The one-to-one mapping between densities and potentials is then established, and we say that the time-dependent potential is a functional of the time-dependent density (and the initial state). This statement implies that, if the only information we have about the system is its density, we can obtain the external potential that produced this density. With the external potential the TDSE can be solved, and all properties of the system obtained. One can conclude that the electronic density determines all other properties of the quantum system. It is important to notice that this is true for a fixed initial state, that is, in addition to the knowledge of n(**r**,t) we also need to know the initial many-body state $|\Psi_0\rangle$.

3.2.2 The Time-Dependent Kohn-Sham equations

In time-dependent systems the total energy is not a conserved quantity, therefore, there can not be variational principle (A.4.1). There exists, however, a quantity analogous to the energy, the *quantum mechanical action* defined as:

$$A[\Psi] = \int_{t_0}^{t} \langle \Psi(t)| i\frac{\partial}{\partial t} - \hat{H}(t)|\Psi(t)\rangle \tag{3.73}$$

where $\Psi(t)$ is an N-body function defined in some convenient space. In their original paper, Runge and Gross [154] offered a derivation of the Kohn-Sham equations starting from the action Eq.(3.73) and constructing an auxiliary system of noninteracting electrons subject to an external local potential ν_{KS}. This potential is unique (by the Runge-Gross theorem applied to the noninteracting system) and is chosen such that the density of the Kohn-Sham electrons is the same as the density of the original interacting system. Thus, total time-dependent electronic density is given by:

$$n(\mathbf{r},t) = \sum_{n=1}^{N} |\psi_i(\mathbf{r},t)|^2 \tag{3.74}$$

Each Kohn-Sham electrons satisfy the following TDSE:

$$i\frac{\partial}{\partial t}|\psi_i(\mathbf{r},t)\rangle = \hat{H}_{HK}(\mathbf{r},t)|\psi_i(\mathbf{r},t)\rangle \tag{3.75}$$

where the time-dependent Kohn-Sham Hamiltonian is defined as:

$$\hat{H}_{HK}(\mathbf{r},t) = -\frac{1}{2}\nabla_i^2 + \nu_{KS}(\mathbf{r},t) \tag{3.76}$$

and the time-dependent Kohn-Sham potential as:

$$\nu_{KS}(\mathbf{r},t) = \nu(\mathbf{r},t) + \nu_J(\mathbf{r},t) + \nu_{XC}(\mathbf{r},t) \tag{3.77}$$

that is, as in the Kohn-Sham scheme for the ground state, the time-dependent Kohn-Sham potential is normally written as the sum of the external, Hartree and exchange-correlation terms. The first one gathers the potential created by the electron-nuclei interaction plus any other time-dependent potential applied. The time-dependent Hartree potential can be written as:

$$\nu_J(\mathbf{r},t) = \int \frac{n(\mathbf{r}_2,t)}{r_{12}} d\mathbf{r}_2 \tag{3.78}$$

Finally last term of Eq.(3.77), the exchange-correlation potential, includes all nontrivial many-body effects, and has an extremely complex (and essentially unknown) functional dependence on the density. This dependence is clearly nonlocal, both in space and in time, i.e., the potential at time t and position $\mathbf{r}$ can depend on the density at all other positions and all previous times. As it happens in ground-state density the results obviously depends on the quality of the approximation. Explicit density functionals, like the adiabatic LDA, only retain the density dependence.

In the original paper of Runge and Gross [154], the exchange-correlation potential was simply the functional derivative. Years after, it was discovered that this formulation suffered from fundamental problems [163]. Using the Keldysh formalism [164] and introducing a new action functional that does not explicitly contain $\partial/\partial t$, a final expression was obtained:

$$\nu_{XC}(\mathbf{r},t) = \left.\frac{\delta A_{XC}}{n(\mathbf{r},\tau)}\right|_{n=n(\mathbf{r},t)} \tag{3.79}$$

where τ is the Keldysh pseudotime.

3.2.3 Adiabatic Approximation

As we have already noticed, the exact exchange-correlation potential depends on the entire history of the density, and also on the initial wavefunctions of both the interacting and non-interacting systems.

The dependence on the initial wavefunction disappears whenever a non-degenerate ground state, for both interacting and noninteracting, is considered: the initial wavefunctions themselves are functionals of the initial density. However, the time and spatial dependences are meaningful since exchange-correlation potential at $\mathbf{r}$ and t depends not just on n($\mathbf{r}$, t) but on all n($\mathbf{r}$,t') for $0 \leq$ t' $\leq$ t, and for arbitrary points $\mathbf{r}$ in space. Thus the potential remembers the density's past, and it is said it has *memory*.

In the adiabatic approximation exchange-correlation potential is approximated as being local in time, that is, all dependence on the past is ignored and the dependence is then reduced to the instantaneous density[4]:

$$\nu_{XC}[n](\mathbf{r},t) \approx \nu_{XC}[n(t)](\mathbf{r}) \tag{3.80}$$

If the time-dependent potential changes very slowly (adiabatically), this approximation will be valid. However, the electrons will remain always in their instantaneous ground state, and in practice, the spatial nonlocality of the functional is also approximated:

$$\nu_{XC}[n](\mathbf{r},t) \approx \nu_{XC}^{GS}[n_{GS}](\mathbf{r})|_{n_{GS}(\mathbf{r'})=n(\mathbf{r'},t)} \tag{3.81}$$

where $\nu_{XC}^{GS}[n_{GS}](\mathbf{r})$ is the exact ground-state exchange-correlation potential of the density $n_{GS}(\mathbf{r})$. Due to the resemblance to the the argument made to determine the function used in LDA calculations for the ground-state energy, the approximation receives the name of *Adiabatic LDA* (ALDA) *approximation*.

3.3 Outline of molecular dynamics simulation based on time propagation

To get know the time-dependent Kohn-Sham wavefunctions at certain time t one can solve the time-dependent Kohn-Sham equations (3.75) or, alternatively, tackle the problem under a different point of view: propagate the electronic orbitals on time in which is called *time propagation scheme*. Numerically, one approximates the time evolution:

$$|\psi(t)_i\rangle = U(t,t_0)|\psi_i(t_0)\rangle \qquad i = 1,...,N_e \tag{3.82}$$

where:

$$U(t,t_0) = \hat{T}exp\left(-\int_{t_0}^{t} H_{KS}(\tau)d\tau\right) \tag{3.83}$$

and $\hat{T}$ is the time ordering operator [165]. The approach we adopt in this thesis to perform the Ehrenfest MD is based on the iterative scheme developed by Baer and Gould [33,166], combined with a two-step Runge-Kutta scheme to maintain order Δt^3 accuracy. The solutions of the time-dependent Schrödinger equations for both half and full steps is accomplished by iterating until convergence the set of integral equations:

$$|\psi_i^{(n)}(t_0+\Delta t)\rangle = |\psi_i^{(0)}(t_0+\Delta t)\rangle - i\int_{t_0}^{t_0+\Delta t} H_{KS}(\{\psi^{(n-1)}(\tau)\},\tau)\psi_i^{(n-1)}(\tau)d\tau \tag{3.84}$$

A general outline can be given by the following steps:

1. In a initial step, the following elements are read:
   - A set of atoms, initial positions and velocities $\{\mathbf{r}_0\}$, $\{\mathbf{v}_0\}$
   - The basis cut-off and the approximation for the exchange-correlation potential.
   - A time-step, Δt, for the atomic movement.

[4]Notation: For clarity we consider the dependence of ν_{XC} on the electron density, $\nu_{XC}[n]$. This statement is also applicable to $\nu_J[n]$

- An initial time-dependent electron density $n_0(\mathbf{r},t)$. (In our special case of double ionizations this density is made by all the KS orbitals of the neutral species except the one removed)

2. Given the initial density, the atomic forces, $\{\mathbf{f}_0\}$, are calculated using Helmann-Feynman theorem (2.31)

3. A first Runge-Kutta step is initiated, the KS orbital are propagated until a middle point using $\nu_{KS}(\mathbf{r}, t_0)$ and Eq.(3.84) to obtain effective potential at time $t = t_0 + \Delta t/2)$. The integrals are resolved by Crank-Nicholson interpolation.

4. In the second step, the full time evolution is then achieved by evolving the wavefunctions for the full time step Δt, using the approximated potential computed from the half step, $\nu_{KS}(\mathbf{r}, t_0 + \Delta t/2)$.

5. With the new density $n_0(\mathbf{r}, t+\Delta t)$, a new set of forces $\{\mathbf{f}\}$ are obtained. The new positions and velocities, $\{\mathbf{r}\}$ and $\{\mathbf{v}\}$, are straightforwardly updated.

6. If the total time of the simulation is reached the simulation is stopped, otherwise, step 3 is reinitiated.

Chapter 4

Observables

Computer simulation allows us to study properties of many-particle systems. However, not all properties can be directly measured in a simulation, and also, most of the quantities given by a simulation do not corerspond to properties that are measured in real experiments. For example, in a MD simulation of liquid water, we could measure the instantaneous positions an velocities of all molecules in the liquid. However, this kind of information cannot be compared to experimental data because no real experiment provides us such detailed information. Rather, a typical experiment measures an average property, averaged over the time of the measurement.

The connection between microscopic simulations and macroscopic properties is made via statistical mechanics which provides the rigorous mathematical expressions that relate macroscopic properties to the distribution and motion of the atoms and molecules of the N-body system; molecular dynamics simulations provide the means to solve the equation of motion of the particles and evaluate these mathematical formulas. In the next sections we will introduce some basics definitions on this field that have been handled in the MD simulation performed in this thesis. A more profund description can be found in references [48, 50, 51, 167–169].

4.1 Ensembles

Statistical mechanics is the branch of physical sciences that studies macroscopic systems from a molecular point of view. The goal is to understand and to predict macroscopic phenomena from the properties of individual molecules. The system could range from a collection of solvent molecules to a solvated protein-DNA complex. In order to connect the macroscopic system to the microscopic system, time independent statistical averages are often introduced. We start this discussion by introducing a few definitions:

- The thermodynamic state of a system is usually defined by a small set of parameters, for example, the temperature, T, the pressure, P, and the number of particles, N. Other thermodynamic properties may be derived from the equations of state and other fundamental thermodynamic equations.
- The mechanical or microscopic state of a system is defined by the atomic positions, $\mathbf{r}$, and momenta, $\mathbf{p}$; these can also be considered as coordinates in a multidimensional space called phase space. For a system of N particles, this space has 6N dimensions. A single point in phase space, denoted by $\mathbf{G}$, describes the state of the system.
- An ensemble generally is defined as a collection of all possible systems which have different microscopic states but have an identical macroscopic or thermodynamic state. In MD

simulations, a sequence of points in the phase space is generated as a function of time. These points represent microscopic states of the system and are impossed to belong to the same ensemble, that is, they must satisfy the conditions of a particular thermodynamic state. There exist different ensembles with different characteristics:

- Microcanonical ensemble (NVE) : The thermodynamic state characterized by a fixed number of atoms, N, a fixed volume, V, and a fixed energy, E. This corresponds to an isolated system and is the one used in the MD performed in our simulations.
- Canonical ensemble (NVT): This is a collection of all systems whose thermodynamic state is characterized by a fixed number of atoms, N, a fixed volume, V, and a fixed temperature, T.
- Grand canonical ensemble (μVT): The thermodynamic state for this ensemble is characterized by a fixed chemical potential, μ, a fixed volume, V, and a fixed temperature, T.

4.2 Time averages

An experiment is usually made on a macroscopic sample that contains an extremely large number of atoms or molecules, sampling an enormous number of conformations. In statistical mechanics, the average values are defined as ensemble averages given by:

$$\langle A \rangle_{ensemble} = \int \int A(\mathbf{p}, \mathbf{r}) \rho(\mathbf{p}, \mathbf{r}) d\mathbf{p} d\mathbf{r} \tag{4.1}$$

where $\langle A \rangle_{ensemble}$ is the observable of interest and it is expressed as a function of the momenta, $\mathbf{p}$, and the positions, $\mathbf{r}$, of the system. The integration is over all possible variables of $\mathbf{r}$ and $\mathbf{p}$.

The probability density of the ensemble is given by:

$$\rho(\mathbf{p}, \mathbf{r}) = \frac{1}{Q} exp\left[-H(\mathbf{p}, \mathbf{r})/k_B T\right] \tag{4.2}$$

where H($\mathbf{p}$,$\mathbf{r}$) is the Hamiltonian of the system, T is the temperature, k_B is Boltzmann's constant and Q is the partition function:

$$Q = \int \int exp\left[-H(\mathbf{p}, \mathbf{r})/k_B T\right] d\mathbf{p} d\mathbf{r} \tag{4.3}$$

This integral is generally extremely difficult to calculate because one must determine all possible states of the system. In a molecular dynamics simulation, the points in the ensemble are calculated sequentially in time, so to get to know an ensemble average, the simulation must pass through all possible states corresponding to the particular thermodynamic constraints. Another way to tackle this problem is to determine a *time average* of A, which is expressed as:

$$\langle A \rangle_{time} = \frac{1}{t} \int_{t_0}^{t_0+t} A(\mathbf{p}(\tau), \mathbf{r}(\tau)) d\tau \tag{4.4}$$

where t is the simulation time[1], S is the number of time steps in the simulation and A($\mathbf{p}$,$\mathbf{r}$) is the instantaneous value of A.

The dilemma appears because one can calculate time averages by molecular dynamics simulation, but the experimental observables are supposed to be ensemble averages. In a practical

[1]Symbol t have been chosen to differentiate a finite time-step, from τ, the magnitude Time.

way, if it is assumed the equilibrium, the interval average of Eq.(4.4) reliably approximates the time average of $\langle A \rangle_{ensemble}$. Thus the later would be obtained from a measurement performed over an infinite duration:

$$\langle A \rangle_{esemble} = \lim_{t \to \infty} \frac{1}{t} \int_{t_0}^{t_0+t} A(\mathbf{p}(\tau), \mathbf{r}(\tau)) d\tau \approx \frac{1}{S} \sum_{t=1}^{S} A(\mathbf{p}, \mathbf{r}) \tag{4.5}$$

Equation 4.5 states one of the most fundamental axioms of statistical mechanics, the ergodic hypothesis. The basic idea is that if one allows the system to evolve in time indefinitely, that system will eventually pass through all possible states. The goal, therefore, of a molecular dynamics simulation is to generate enough representative conformations such that this equality is satisfied. If this is the case, experimentally relevant information may then be calculated if one is certain to sample a sufficient amount of phase space.

4.3 Temperature in Molecular Dynamics

One situation in which the ergodic assumption is obviously valid is when the integrand A(**p**,**r**) is a constant of the motion. In that case A does not vary along the phase-space trajectory and any interval average given by Eq.(4.4) is equal to the time average of Eq.(4.5). However as the phase point journeys along the hypersurface of constant energy E, most quantities are not constant: instead their values fluctuate. That is because the molecules are in continual motion and are colling with one another, the positions and momenta of individual molecules are continually changing and therefore most functions that depend on the positions and momenta fluctuate.

One example of a fluctuating quantity is the translational kinetic energy, whose average tends to be stable about a value proportional to the kinetic temperature:

$$\langle E_k \rangle = \lim_{t \to \infty} \frac{1}{t} \int_{t_0}^{t_0+t} E_k(\mathbf{p}(\tau)) d\tau = \frac{k_B T}{2}(3N - N_c) \tag{4.6}$$

where:

$$E_k = \sum_{i=1}^{n} \frac{1}{2m_i} \mathbf{p}_i^2 \tag{4.7}$$

and N is the degrees of freedom, N_c the number of constraints, n the total number of particles and m_i is the mass of the i-particle.

Sometimes, it is interesting to check the effects that the increasing or decreasing of temperature generates in our simulated system. In other cases, such as phase transitions, the control has to be rigorous since this variable has to remain as a constant. Many schemes have been developed in this sense, starting from the scale of the velocities $\{\mathbf{v}\}_i$ [170], in which the temperature of the system at certain time T(t) is multiplied by a factor λ. The associated change can be calculated as follows:

$$\Delta T = \frac{1}{2} \sum_{i=1}^{N} \frac{2m_i(\lambda \mathbf{v}_i)^2}{3Nk_B} - \frac{1}{2} \sum_{i=1}^{N} \frac{2m_i \mathbf{v}_i^2}{3Nk_B} = (\lambda^2 - 1)T(t) \tag{4.8}$$

$$\lambda = \sqrt{T_{new}/T(t)} \tag{4.9}$$

notice that the number of constraints, N_c, have been reduced to zero. The simplest way to control the temperature is thus to multiply the velocities at each time step by the factor $\lambda = \sqrt{T_{req}/T_{curr}}$

where T_{curr} is the current temperature as calculated from the kinetic energy and T_{req} is the desire temperature.

An alternative way to maintain the temperature is to couple the system to an external heat bath that is fixed at the desired temperature [171]. The bath acts as a source of thermal energy, supplying or removing heat from the system as appropriate. The velocities are scaled at each step, in such a way that the rate of change of temperature is proportional to the difference in temperature between the bath and the system. Extended system methods, originally introduced by Nosé [172] and subsequently developed by Hoover [173], consider the thermal reservoir to be an additional degree of freedom with certain potential energy a fictitious mass.

4.4 Charge analysis on DFT calculations

In spite of the special relevance that the atomic charges have for chemists, assigning charges to a molecule in a cluster, or to an atom into a molecule is not a trivial task. The atomic charges are not observable and therefore, not defined by quantum mechanical theory. Once we get, for instance, a converged electronic density from a quantum calculation it is not clear how we should partition it amongst the fragments of the system. Many different partitioning schemes have been proposed and they can roughly be divided into two main categories: orbital-based methods, and methods based on electrostatic potential or electron density.

In the former category, the earliest of such schemes used in *ab-initio* framework was the Mulliken population analysis [174, 175]. This method partitions, in a somewhat arbitrary way, the contribution of each occupied spin-orbital to the total population of each basis function, and also the contribution to the overlap population between each pair. There are others population analysis based on electronic orbitals, for instance, the study of the *Natural Atomic Orbitals* suggested by Weinhold [176] based on the previous work of Löwdin on the *Natural Orbitals* [177]. However, it is worth noting that atomic basis sets are overcomplete and such decomposition is not unique. In principle, a calculation could be done where all the basis functions are located on one of the atoms in the system, which would then lead to an assignment of all the electrons in the system to that atom.

One less arbitrary fashion of approaching this problem is the decomposition of the charge density proposed by the Richard Bader's theory of Atoms in Molecules (AIM) [178, 179] based on the electron density. This theory has been demonstrated to be very adequate with methods using DFT and specially with those using plane wave basis functions that are not associated with any particular atom in the system.

Within this context, each point of the space is assigned to one subsystem separated by surfaces in which *gradient vector field of the electro density* $\bigtriangledown n(\mathbf{r})$ is orthogonal to the *normal vector of the density surface* $z(\mathbf{r})$ (they do not have normal component of the density gradient). These normal *zero-flux surfaces* $S(\Omega, \mathbf{r})$ are mathematically defined by:

$$\bigtriangledown n(\mathbf{r}) z(\mathbf{r}) = 0 \qquad \forall \mathbf{r} \in S(\Omega, \mathbf{r}) \tag{4.10}$$

The dominant topological property of a molecular charge distribution is that it exhibits maxima at the positions of the nuclei, therefore these surfaces can be considered as a natural consequence of having maximums together.

The boundary condition imposed by Bader leads to the partitioning of a molecular system into a set of disjoint spatial regions, each region containing a maximum (although is not necessary [180]) that corresponds to a single atom. The electronic population of a particular atom A, is

thus given by:

$$N_A = \int_{\Omega_A} n(\mathbf{r}) d\mathbf{r} \tag{4.11}$$

It can be also shown that the atomic subsystems obey a local virial relation. Moreover, other properties further than atomic charge, like atomic dipole moment, atomic kinetic or potential energies, are shown to be well-defined by quantum mechanics if they are also integrated over these atomic subvolumes [181]. The properties of the gradient vector field also contain the information needed for a definition of molecular structure and its stability by using the mathematics of qualitative dynamics [182, 183]. So that, a topological study of the density surface by means of the gradient density $\bigtriangledown n(\mathbf{r})$ indeed reveals a lot of information about the system.

The mapping of this vector is done by, first, localizing the *Critical Points* (local maxima, minima and saddle points) achieved by the calculation of its roots, and second, the characterization of these points via associated Hessian matrix **H**:

$$\mathbf{H} = \begin{pmatrix} \frac{\partial n(\mathbf{r})}{\partial^2 x} & \frac{\partial n(\mathbf{r})}{\partial x \partial y} & \frac{\partial n(\mathbf{r})}{\partial x \partial z} \\ \frac{\partial n(\mathbf{r})}{\partial y \partial x} & \frac{\partial n(\mathbf{r})}{\partial^2 y} & \frac{\partial n(\mathbf{r})}{\partial y \partial z} \\ \frac{\partial n(\mathbf{r})}{\partial z \partial x} & \frac{\partial n(\mathbf{r})}{\partial z \partial y} & \frac{\partial n(\mathbf{r})}{\partial^2 x} \end{pmatrix} \tag{4.12}$$

Diagonalization of **H** in the critical points of the density $n(r_c)$ gives the eigenvalues h_{xx}, h_{yy} and h_{zz} and with this knowledge the different points can be characterized by their "rank" and "signature", symbolically written as "(*r*,*s*)". The rank r is defined as the number of non-zero eigenvalues of the Hessian and the signature s is the sign of the determinant evaluated at the location of the critical point.

The result is a theory of atoms, bonds, structure, and structural stability, where the method does not require the assignment of each basis function to a specific atom. This makes AIM particularly advantageous in situations where this assignment would be difficult or impossible, such as non-localized basis sets.

4.4.1 Finding the zero flux surfaces

There have been proposed several methods to analyze the atoms charges following the topologycal definition of Bader. The idea can be view as a "simple" two-step process in which, in some manner, the Bader volumes are defined and after the electron density $n(\mathbf{r})$ is integrated within each rigion. Unfortunately, one of the biggest problems of the AIM theory is the fast and accurate calculation of these regions or equivalently their non-zero surfaces. It is indeed a mathematical challenge, specially in systems with sudden changes in the density surface.

Attending to how the surfaces are calculated one can find different techniques. For instance, in the early algorithms implemented for small molecules the gradient of the charge density is calculated from derivatives of an analytic wavefunction [178,184]. These methods find stationary points in the charge density and after follow trajectories along the density gradient from these points to map out their connectivity and the zero-flux dividing surfaces as Fig.(4.4.1) shows. With the dividing surfaces represented in this way, the charge in each Bader volume can be integrated radially from the charge density maximum to the surface. However, while this approach works for small molecules, a high density of descent trajectories is needed to accurately represent the surface away from the critical points. For this reason, the method results computationally expensive for large systems and has been harshly criticized, giving rise to new approaches.

To perform the charge calculations of this thesis we have used a free software program (available in *http://theory.cm.utexas.edu*) based on the *fast and robust algorithm* of Henkelman

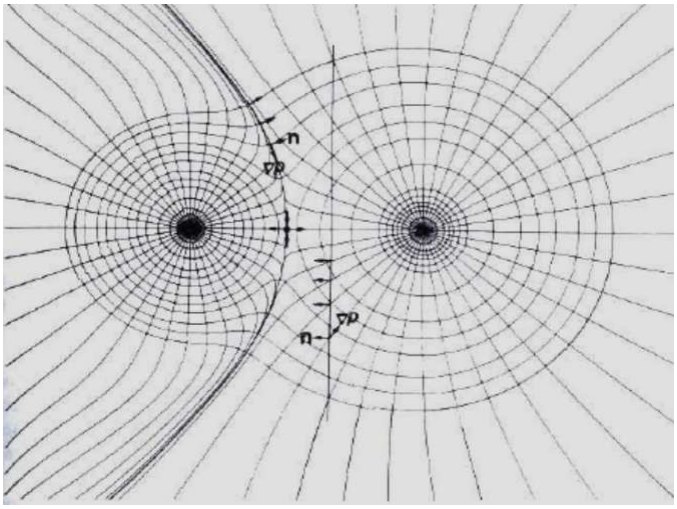


Figure 4.1: Mapping of $\bigtriangledown n(\mathbf{r})$ between two atoms. The isodensities sourround the atoms while the calculated trajectories are the "radial" lines that start in the nuclei and go away. The non-zero surface is represented by the bold line

et al. [185]. In next pages we will describe the method, in its first approach, and we will also comment the biggest improvement [186] of the latest *Version 0.25c*.

4.4.2 Henkelman's grid-based algorithm

Typically, the output of DFT-based calculation is a continuous electronic charge density function $n(\mathbf{r})$ stored in a file a three-dimensional regular Cartesian mesh of values defined over a rectangular volume of space V. The size of this volume should be large enough so that the value of $n(\mathbf{r})$ is negligible outside this region. The original grid-based Bader analysis presented here [185] takes this density file as starting point and subdivides the volume V into M smaller volumes δV_i, $i = 1, 2...M$ centered around each mesh point.

Rather than explicitly representing the interatomic surfaces, the spirit of the algorithm is to consider the problem of finding the Bader volumes under a different point of view, i.e., each volume element δV_i is assigned to a Bader region by tracing a trajectory from the center of the volume to an atomic nucleus (a maximum). What is followed during the trajectory is the direction that maximizes an approximated charge density gradient projected over the direction that links the initial grid point to one of its 26 neighbors. In this sense, all the grid points that make up the trace to the maximum are assigned to that atom. Once the whole assignation is accomplish, the total charge is given by integrating the charge density over the grid points of that region. Since the trajectory true density gradient is contrained to the mesh directions this method is also referred as "on-grid" algorithm.

The Fig.(4.4.2) has been taken from Ref. [186] and shows a visual idea of the process. On the left picture some ascent paths are shown. As an example, starting from the first point on the left of the upper row, a maximum is reached in a four step trajectory. After, a second point is chosen, for instance, the second one of the first row. In this case the trajectory leads to an already assigned point in two steps. Then, a third point is chosen and so on. As it can be seen all the points converge in two different maximums that are separated by a zero flux surface represented by the red line in the figure on the right.

A more detailed view of the procedure brakes it down into the following steps:

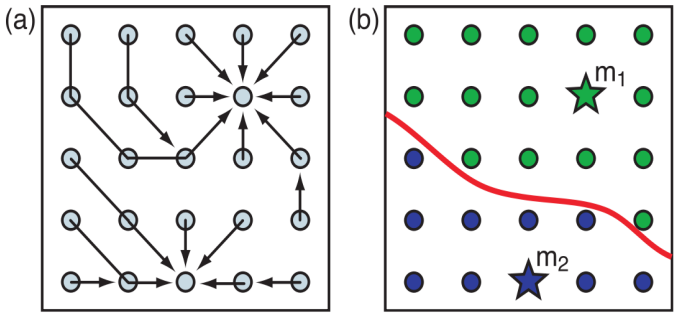


Figure 4.2: Drawing a), illustration of the steepest ascent paths traced by the Henkelman's grid-based algorithm. After all grid points are assigned (b), the set of points which terminate at each maximum (green to m_1 and blue to m_2) constitute that Bader volume. The Bader surfaces (red) separate the volumes.

(i) First an initial grid point is chosen, located by a position vector $\mathbf{r} = l\Delta x\mathbf{i} + m\Delta y\mathbf{j} + n\Delta z\mathbf{k}$; where l, m, and n are integers that indicate the discrete numbering of the mesh points and $\Delta x, \Delta y$ and Δz are the mesh spacings along each axis.

(ii) The projection of the charge density gradient along the direction to each of the 26 neighboring grid points is approximated using finite differences:

$$n_{\Delta\mathbf{r}} \equiv \frac{n(\mathbf{r}+\Delta\mathbf{r}) - n(\mathbf{r})}{|\Delta\mathbf{r}|} \tag{4.13}$$

Where $\Delta\mathbf{r}$ is the position vector to the neighbor reached by the grid step $(di, dj.dk)$. The integers di, dj, and dk can each take the values 1, 0, 1 (excluding $di = dj = dk = 0$). Notice that expression (4.13) is not a vector, but a scalar. It means an approximated value of the gradient density projection over the specified direction, but not a projection vector in the mathematical sense.

(iii) The neighbor which posses the maximum projection from Eq.(4.13) is determined as the next point along the ascent path. If the point has been already assigned to a maximum then the trajectory is stopped and step (i) is re-initiated. Otherwise, the current position vector is stored in an array.

(iv) If non of the projections are positive, the current mesh point is identified as a maximum. All the stored positions that contitute the trace to this point are attributed to it.

(v) The process is repeated until all the volumes are assigned to a maximum.

We must insist that the algorithm does not calculate such surfaces, only their location can be infer after the assignment of grid points. The method is highly efficient, scales linearly with system size, and is also robust to complex bonding topology found in condensed systems. However, it introduces severe lattice bias specially when the true dividing surface runs at a small angle to the grid over long distances.

4.4.3 Fixing the lattice bias

The partialities on the "on-grid method" were firstly pointed out by the work of Sanville et al [187]. The gradient can follow any of the infinite directions of the whole space, however, in the algorithm its projection is constrained to only 26 of them. As a consequence, the trajectories accumulate error whenever the gradients lines do not run parallel to the lattice directions and it

leads to wrong assignations, particularly grave near of Bader surfaces. In these zones the gradient barely changes and one mistake done in a grid point is easily translated to next one. These bias can be solved increasing the grid points until each projection finds the correct direction to project, circumstance that unfortunately, occurs only in a ideal infinite grid.

Recently, W. Tang and et [186] have worked out the problem by introducing a correction vector $\Delta\mathbf{r}$ that adjusts the trajectories. The vector is defined as:

$$\Delta\mathbf{r} = \Delta\mathbf{r} + (\mathbf{r}_{grad} - \mathbf{r}_{grid}) \tag{4.14}$$

and takes into account the difference between the direction of the gradient vector, $\mathbf{r}_{grad}$, and the direction of the step just performed, $\mathbf{r}_{grid}$. The gradient vector is given by the direction that maximizes the change in the gradient density:

$$\mathbf{r}_{grad} = c(\nabla n_x, \nabla n_y, \nabla n_z) \tag{4.15}$$

evaluated at the six closet neighbors using a central finite difference:

$$\begin{aligned}
\nabla n_x &= \frac{n(i+1,j,k) - n(i-1,j,k)}{\mathbf{r}(i+1,j,k) - \mathbf{r}(i-1,j,k)} \\
\nabla n_y &= \frac{n(i,j+1,k) - n(i,j-1,k)}{\mathbf{r}(i,j+1,k) - \mathbf{r}(i,j-1,k)} \\
\nabla n_z &= \frac{n(i,j,k+1) - n(i,j-1,k)}{\mathbf{r}(i,j,k-1) - \mathbf{r}(i,j,k-1)}
\end{aligned} \tag{4.16}$$

with the constant equal to:

$$c = min(dx/|\nabla n_x|, dy/|\nabla n_y|, dz/|\nabla n_z|) \tag{4.17}$$

The algorithm is particularly well suited to DFT calculations of large molecules or materials such as surfaces, alloys or molecules in solution [188–190]. It can be resume in the following steps:

(i) Starting from an initial point, the charge density gradient is projected over the six closest neighbors following Eq.(4.16).

(ii) The gradient vector $\mathbf{r}_{grad}$ is then calculated according to Eq.(4.17). This vectors points in the direction of the gradient density and normally runs off the grid. Since we desire to retain the trajectory on the grid, a jump to the nearest point is done. The step between them is labeled by the vector $\mathbf{r}_{grid}$ and the deviation from the true trajectory is calculated by means of the correction vector of Eq.(4.14) (initially zero).

(iii) When the length of any component of $\Delta\mathbf{r}$ is larger than half of the grid spacing, a correction step is taken in that direction. The correction vector is then recalculated by subtracting the correction step. In this way, the true trajectory is never more than half a grid point from the current grid point in any direction.

(iv) The ascent trajectory is ended if a charge density maximum is reached or, at some point, there is no non-assigned neighbors.

Depending upon the order in which grid points are analyzed, the grid point adjacent to the Bader surface can be assigned to one of the two volumes on either side of the dividing surface. This ambiguity is due to the fact that the trajectory between grid points deviates from the true trajectory by up to half a grid step. Thus, when all the points are assigned, a final refinement of the grid points is required to identify all grid points on the boundary of a volume.

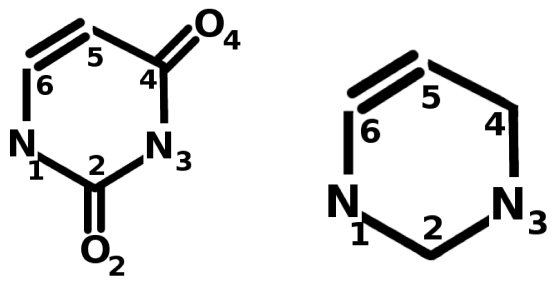
Part III

Results

Introduction

In the present Chapter we will summarize the results obtained by applying the combination of TD-DFT MD and BO MD to the study of double ionizations of different biomolecules. The writting have been divided attending to the systems investigated:

- In Section 5: We will start by describing the fragmentation in gas phase of uracil^{2+}, comparing the results with the experimental data obtained by the proton collision experiment of Moretto-Capelle et al in Toulouse (France) [191].
- In Section 6: The theoretical study of uracil^{2+} in liquid phase will be presented. The comparison with gas phase results will allow us to estimate how the presence of water affects the fragmentation processes. Besides, we will analyze the effects of the ionization when it is produced in a surrounding water molecule near the uracil.
- In Section 7: We will investigate the double ionizations of small water clusters compared with the experimental data obtained by the COLTRIMS experiment [192] carried out in Caen (France). As we will see our theoretical calculations are able to describe the fast mobility occurring in the clusters after the collision.
- In Section 8: Preliminary results of double ionizations of pyrimidine will be presented.



(a) Uracil (b) Pyrimidine

Figure 4.3: Labels used for uracil and pyrimidine

The labels used for atoms of the systems studied are displayed in Fig.(4.3). All dynamics have been run using the Car-Parrinello Molecular Dynamics code CPMD 3.13.2 [152] with the special TD-DFT MD Ehrenfest implementation of Tavernelli et al [33]. Other conditions such as supercell size, time-step or cutoff basis set will be given for each particular case. We have also carried out all-electron calculations using Gaussian09 package [193] in order to support conclusions. In these cases BLYP exchange-correlation functional [116, 119] and standard Pople basis sets [194–198] have been used.

Chapter 5

Fragmentation of Uracil^{2+} in gas phase

Abstract *We apply a combination of Time-Dependent Density Functional theory and Born-Oppenheimer molecular dynamics methods to investigate fragmentation of doubly-charged gas-phase uracil produced in collisions with 100 keV protons. The results are in good agreement with ion/ion coincidence measurements performed for the occasion. Fragmentation mainly arises from inner shell two-electron vacancies. Orbitals of similar energy and/or localized in similar bonds can lead to very different fragmentation patterns, thus showing the importance of the intramolecular chemical environment. The observed fragments almost never correspond to the energetically most favorable dissociation paths, which is due to dynamical effects occurring in the first few femtoseconds after electron removal*

5.1 Introduction

To unravel the mechanisms at this early stage after the radiation of biological tissues by highly charged ions, physicists have performed numerous experiments in gas phase (see e.g. [35,36,199]), in which swift charged ionic projectiles impinge on DNA/RNA bases, sugars, nucleotides or even biomolecular clusters. In contrast with experiments performed in solution or directly on living systems, gas phase experiments provide direct and precise information on *single* collision events. Thus they allow one to unambiguously identify fragmentation channels associated with a given biomolecule and not with the environment. Such detailed information can be achieved by combining techniques that are state-of-the-art in gas phase chemistry, e.g., high resolution mass spectrometry, and in collision physics, such as multicoincidence detection techniques that provide the correlation between different charged fragments as well as their relative kinetic energies and momenta.

Fragmentation results from relaxation of the excess electronic energy associated with vacancies created in the different electronic shells of the molecule. Since, in these collisions, electrons can be removed from many of these shells, experiments cannot tell us how fragmentation depends (i) on the shape/energy of the molecular orbital (MO) in which the electron vacancies are created, and (ii) on the intramolecular environment, i.e., on the neighboring functional groups. This information can only be obtained from *ab-initio* MD calculations such as those based on TD-DFT methods [154], which have been already applied to investigate the radiolysis of bulk water [34] and water clusters [192]. In the group, we apply TD-DFT MD and report new experi-

mental results to investigate the fragmentation of doubly ionized uracil created in collisions with 100 keV protons. We show not only that theoretical results are compatible with the experimental ones, but also that fragmentation depends critically on the energy, shape and intramolecular chemical environment of the ionized MO. This suggests that, by targeting specific MOs, damage can be induced rather selectively.

The choice of uracil (an RNA basis) follows a rich tradition in experimental collision physics [24, 25, 35–39]. Uracil was the first biomolecule to be investigated in this context and is often considered as a benchmark due to its simplicity and high stability in gas phase (which prevents, e.g., its thermal degradation). The study of uracil in gas phase is also important to understand the early presence of RNA bases on Earth (e.g., the fact that uracil can be formed in outer space suggests an exogenous origin of life [200]). The most detailed experiments on uracil have been performed with proton projectiles of 20-150 keV [24, 36, 37]. In particular, Tabet et al [36] have reported relative cross sections for electron capture and direct ionization, and the corresponding fragmentation branching ratios. Therefore, H^+/uracil collisions are ideal (i) to check the accuracy of existing theoretical methods and (ii) to obtain, with the help of the latter, dynamical information that experiments alone cannot provide.

To theoretically describe the fragmentation of doubly charged uracil, we take advantage of the fact that the collision is very fast and, therefore, that electron removal from the neutral molecule is sudden compared to the fragmentation time. Thus, one can safely assume that the geometry of uracil^{2+} and the velocities of its nuclei just after the collision are the same as for neutral uracil. We consider two-electron removal from inner orbitals, namely, the lowest ones obtained in the Kohn-Sham pseudo-potential representation of neutral uracil described below (9a, 10a and 11a, hereafter called KS1, KS2 and KS3), and from the valence highest occupied molecular orbital 5a (HOMO). The corresponding energies of KS1, KS2, KS3 and HOMO orbitals are -27 eV, -26 eV, -24 eV and -6 eV respectively. The drawings of the orbitals are given by Fig.(5.1) by a red mesh isosurface of 0.018 isovalue. KS1 and KS2 orbitals are associated with localized σ-like C_2-O and C_4-O bonds in a different chemical environment: while KS1 lies between two N atoms, KS2 lies between N and C atoms. KS3 is delocalized over one of the CO bonds and part of the uracil ring. The HOMO is even more delocalized since it extends over half of the ring and the two O atoms bonded to it.

This orbital sampling allows us to investigate not only the influence of the energy and shape of the ionized MO on the fragmentation processes, but also the effect of the chemical environment on a particular bond. To properly account for non adiabatic effects in the fragmentation arising from inner-shell vacancies, we have performed first-principles calculations within the so-called Ehrenfest MD approximation where the mean field potential energy surface driving the nuclear dynamics, Eq.(2.71), is computed at the TD-DFT level.

Due to the very small time step (of the order of attoseconds) used in the real time propagation of electrons, the total simulation time is necessarily limited to less than 100 fs. Since the complete dissociation of the ionized molecule requires in general longer times (a few ps), we then switch to ground-state BO MD when this dynamics becomes essentially adiabatic (i.e., when Ehrenfest and BO MD produce the same trajectory).

Fragmentation associated with HOMO vacancies is studied by using standard Car-Parrinello MD from the very first instant after ionization and thermal equilibration by velocity scaling (in this case, fragmentation is so slow that TD-DFT MD is prohibitively expensive). A detailed description of our TD-DFT MD method can be found in Ref. [33]. Briefly, we use a cubic box of size L=19 Å and a plane wave basis with an energy cutoff of 70 Ry. Core electrons are replaced by pseudopotentials of the standard Troullier-Martins form [144]. The exchange correlation energy is calculated using the GGA functional BLYP [116, 119]. In TD-DFT-based Ehrenfest MD, the

electronic density evolves according to the time-dependent Kohn-Sham (TD-KS) equations [154].

The propagation of the electronic degrees of freedom is started from a non-equilibrium electronic structure in which two electrons are removed from a previously occupied KS orbital of neutral uracil[1].

This initial electron density, which does not correspond to a specific state of uracil^{2+}, is then propagated by numerical integration of the TD-KS equations using an iterative Crank-Nicholson algorithm with a time step δt=0.00024 fs combined with a two-step Runge-Kutta scheme to maintain δt^3 order accuracy. The small step chosen assures energy conservation within 0.02% (0.4 eV). In the BO MD calculations we use δt=0.024 fs. The initial atomic configuration is taken from an equilibrated trajectory obtained from a CP MD simulation of the uncharged system at 350 K. The forces on the nuclei are computed according to the Hellman-Feynman theorem as the analytic derivative of the expectation value of the energy. Atomic charges along the different trajectories are evaluated using Bader's AIM approach as implemented in [201].

For the discussion of the results, we will present first the TD-DFT MD simulations corresponding to KS1, KS2 and KS3 ionization of uracil. Afterwards, we will summarize the data obtained while the dynamics have been extended to longer times using BO MD. At this point, we will also comment HOMO ionization since it has performed using the same ground-state techniques. We will end the section by comparing the theoretical results with the experimental ones obtained in proton collision experiments.

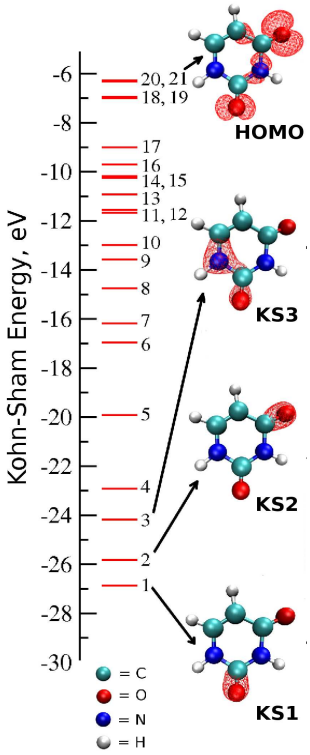


Figure 5.1: Energies of the occupied KS orbitals of neutral uracil and electron densities associated with the KS1, KS2, KS3 and HOMO orbitals (0.018 isovalue).

5.2 TD-DFT dynamics

Fig.(5.2) displays the temperature and the Kohn-Sham energy for KS1, KS2 and KS3 TD-DFT MD. As it can be seen, KS1 and KS2 ionizations represent a much more energetic processes. It is worth to mention that in both cases the Kohn-Sham potential undergoes a change of 0.12 and 0.24 a.u. (3.35 and 6.63 eV) in the first 4 fs of the dynamics. These sudden variations really require the small time-step of integration used, otherwise the total energy of the system would not be conserved and the dynamics might be biased.

Regarding the temperatures of the atoms, they increase due to the exchange of potential energy into kinetic energy during the first fs of dynamic. Particularly, KS2 dynamics has a maximum peak of 5500 K at about 4 fs that, as we will see later, is related with the oxygen loss.

The three fragmentations are schematically summarized in Figures (5.3), (5.4) and (5.5). For all of them, first frame is taken at the beginning of the dynamics with the ionized orbital represented by the red mesh. Main distances are indicated in black and mass and charge of

[1]The electron removal from different MOs is not considered because spin-polarized calculations are prohibitive

fragments are given in blue.

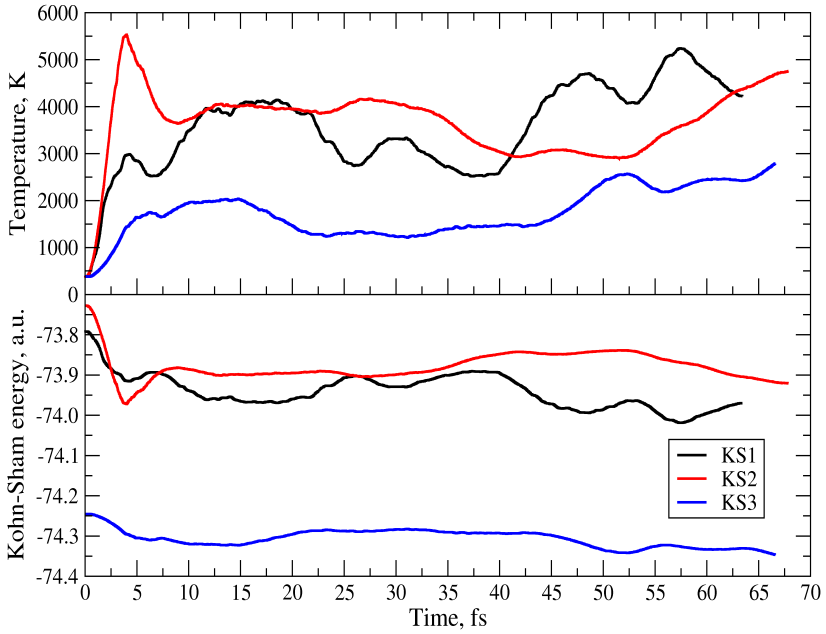


Figure 5.2: Kinetic energy of the nuclei and Kohn-Sham energy of the system during the KS1 (black), KS2 (red) and KS3 (blue) TD-DFT MD

As it is shown, fragmentations are different even for shape-equivalent orbitals like KS1 and KS2. Dynamics of the deepest orbital ionization steers to three fragments of masses 28/41/43 with charges 0.3/1.1/0.6. In this case, the uracil^{2+} ion undergoes the first bond breaking of N_1-C_2 at 14 fs. Bond C_4-C_5 is dissociated at $\approx$ 32 fs, and soon after, C_2-N_3 also breaks ($\approx$ 34 fs). At 35 fs the three fragments are formed and they remain unaltered until the end of the dynamics. As can be seen, the original CO bond from which the electrons were removed is not broken.

Concerning the charges, fragment 41 preserves a charge of 1+ from almost the beginning. Fractional charges of 28 and 43 fragments, however, must be taken with some caution. As it is well known, DFT methods can lead to unphysical fractional charges. A paradigmatic example is H_2^+, for which TD-DFT predicts dissociation into $H^{+0.5}$ + $H^{+0.5}$, instead of into either H + H^+ or H^+ + H. Thus, in general, calculated fractional charges of the order of +0.5 are an indication of the existence of two dissociation paths in which the charge is asymmetrically and alternatively distributed between two fragments. That suggests the formation of a singly charged fragment with mass 28 and a neutral fragment with mass 43, and the other way around.

KS2 dynamics represents the most fragmented case since it leads to two oxygen atoms and three fragments of 26/27/27 masses. The C_4-O bond is broken at 4 fs and the O4 atom is consequently detached with a charge that fluctuates between -0.1 and +0.3 during the trajectory. At 24 fs, fragment 26 can be considered formed since the bonds N_1-C_6 and C_4-C_5 are broken, and it remains $\approx$ +0.5 charged. The fragment that stays behind losses the O2 at 28 fs and remains neutral (as it happens with the other released oxygen). Finally the big fragment 54 is broken leading to two equivalent fragments 27/27 sharing equally the charge.

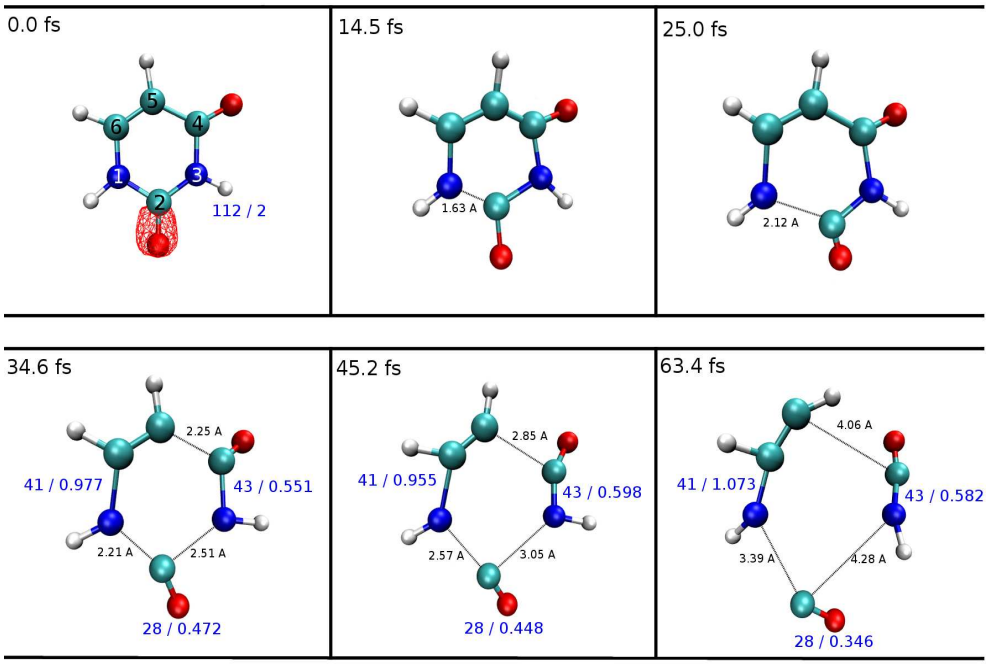


Figure 5.3: Snapshots of the TD-DFT MD trajectory after the KS1 ionization

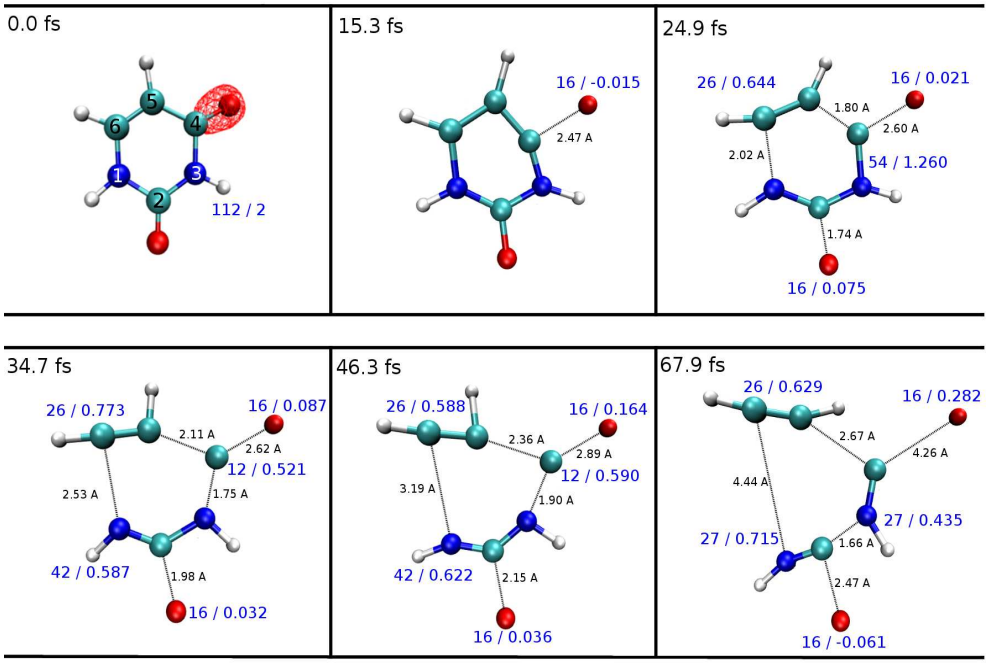


Figure 5.4: Snapshots of the TD-DFT MD trajectory after the KS2 ionization

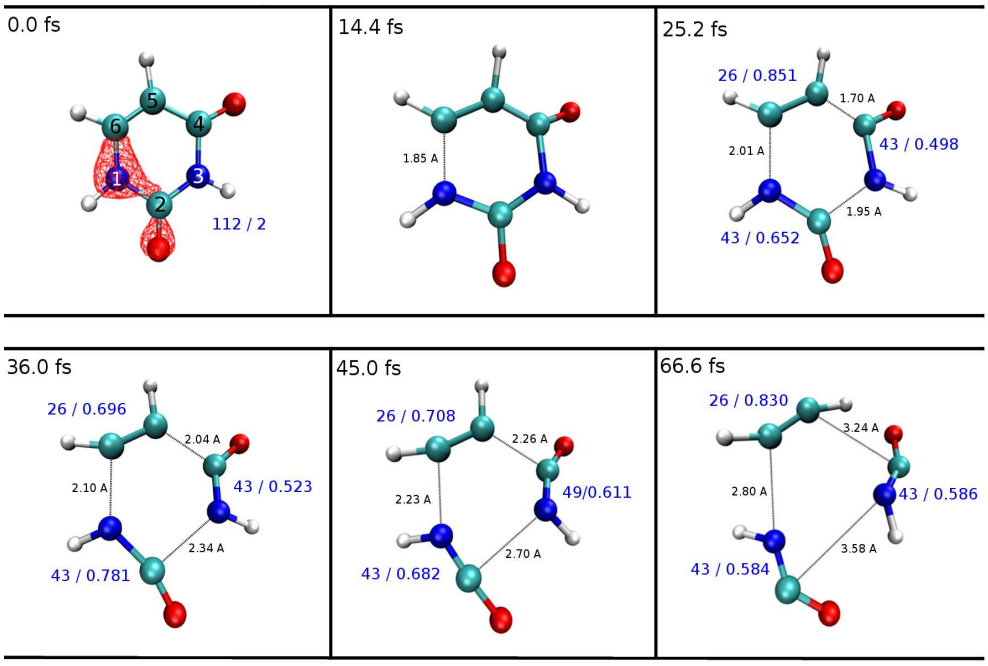


Figure 5.5: Snapshots of the TD-DFT MD trajectory after the KS3 ionization

Ionization of KS3 produces one fragment of 26 mass and two equivalent fragments of 43. First bond breaking of N_1-C_6 occurs at 13 fs while C_4-C_5 and N_1-C_6 break almost simultaneously at about 22 fs. The charge +2 is distributed amongst the fragments in fractional charges.

5.3 Switching to a ground-state surface

In our approach, time-dependent simulations are performed approximately during the first 60 fs that follow the ionization of the target molecule. After, a ground-state dynamics is introduced restarting from last configuration (positions and velocities) coming from the TD-DFT MD outcome. The switching from one method to the other is done once the system evolve with the same trajectory and charges under both schemes.

The results of the theoretical calculations at the end of the TD-DFT MD + BO MD simulation are summarized in Fig.(5.6) and also Table (5.1) that gathers the charges, masses and kinetic energies of the fragments obtained.

During the Born-Oppenheimer dynamics of KS1 orbital the fragment of mass 41 has a charge of about +1 during the trajectory (as it happend during the TD-DFT MD simulation), nevertheless, the atom H6 of C_6 is placed on C_5. Optimizations using all-electron calculations revealed that fragment 41 charged 1+ is stable only if H6 is bounded to C_5, otherwise it does not represent a minimum. The rest of the charge is shared between fragments of mass 28 and 43 that have charges close to +0.5. Thus, KS1 fragmentation is compatible with the 28/41 and 41/43 coincidences commented in next section.

In the case of KS2, one of fragments of 27 mass generated during the TD-DFT MD simulation rotates in such a way that C_4 faces the second oxygen atom detached (O2). In consequence O2 is finally bounded to C_4. Fragments with masses 26, 27 and 43, and charge of the order of +0.5 are produced. None of these fragments are seen in KS1 fragmentation, showing that chemical environment plays a crucial role. Thus, results for KS2 are compatible with coincidence signals at 26/27, 26/43 and 27/43. On the other hand, for KS3 dynamics, former TD-DFT MD fragments

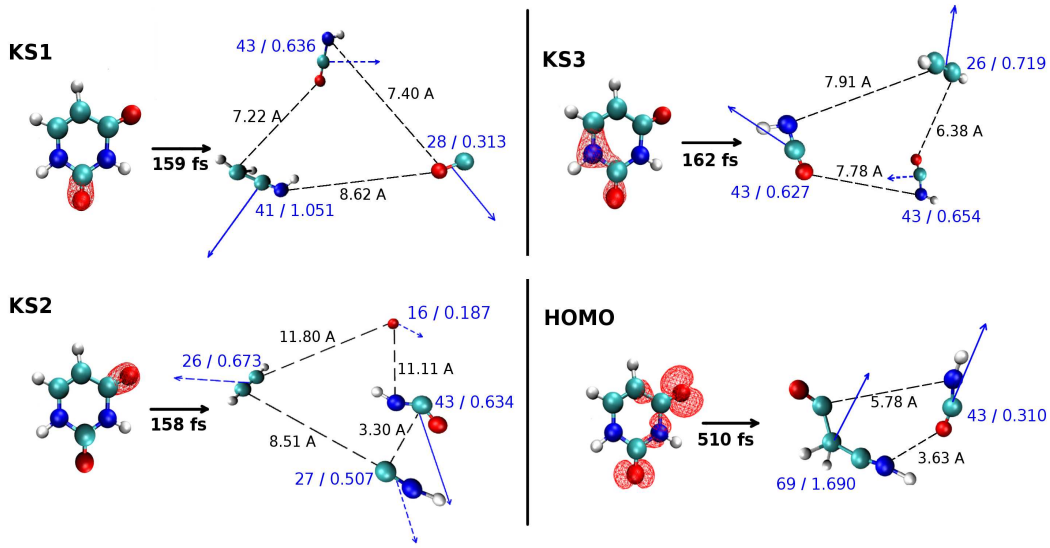


Figure 5.6: Fragments observed at the end of the TD-DFT MD + BO MD propagation of KS1, KS2, KS3 and HOMO (2800K) dynamics. Their corresponding M/Z ratios and center-of-mass velocities (dashed arrows point inwards); fragment distances (dashed lines) are given in Å.

are maintained: a charged fragment with mass 26 plus two charged fragments with mass 43, that lead to a 26/43 coincidence event.

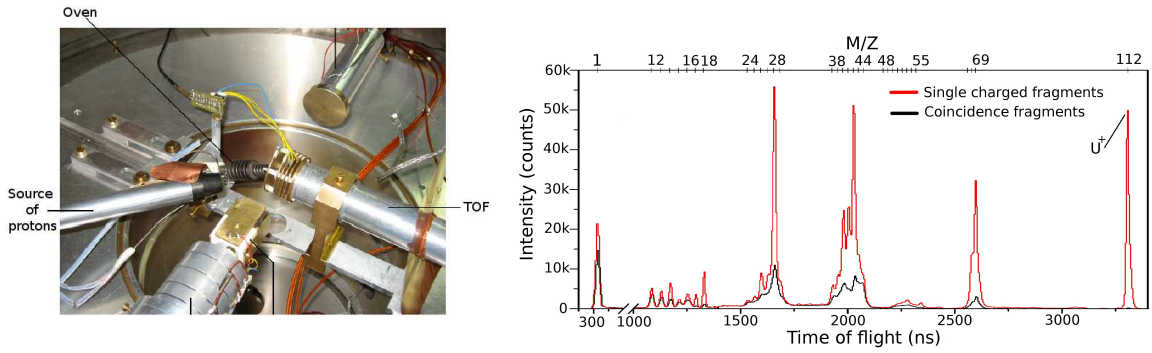
Table 5.1: Mass, charge and kinetic energy of the fragments obtained at the end of BO MD.

Orbital	Fragment	Mass	Charge	Ekin (eV)
$KS1$	CO	28	0.313	1.836
	$HNCCH_2$	41	1.051	1.526
	OCNH	43	0.636	1.284
$KS2$	O	16	0.187	2.011
	HCCH	26	0.673	1.997
	CNH	27	0.507	0.856
	OCNH	43	0.634	0.601
$KS3$	HCCH	26	0.719	1.385
	OCNH	43	0.654	1.228
	OCNH	43	0.627	1.591
$HOMO$	$HNCCH_2CO$	69	1.690	0.214
	OCNH	43	0.310	0.309

HOMO fragmentation have been performed by standard CP MD since the fragmentation takes place in the ground-state of uracil^{2+} and, consequently, takes much longer time. In fact, no fragmentation is observed in CP MD simulations at 350 K extended up to 5 ps. For this reason, we have also performed calculations at 2300 and 2800 K, so that the initial average kinetic energy per atom is comparable to the energy available when two electrons are removed from an inner shell. In this case, fragmentation occurs in $\approx$ 0.5 ps. The dissociation products are the same for both temperatures leading to a 43/69 coincidence.

The experimental setup used in this work has been described in detail in Ref. [202]. Briefly, a pulsed proton beam of 100 keV and 2 mm diameter collides an effusive uracil gas jet produced by heating commercial powder purchased at Sigma-Aldrich at a temperature of 120-150 °C. Charged

fragments produced in the collision are mass-over-charge (M/Z) analyzed by means of a time-of-flight mass spectrometer operating in second order space focusing with pulsed extraction.



(a) Experimental set-up

(b) Time-of-flight spectrum

Figure 5.7: On the left part (a), picture of the experimental set-up in which the oven, source jet and TOF detector are pointed out. On the right side (b), time-of-flight spectr of singly charged uracil fragments (red line) compared to that of charged fragments measured in coincidence (black line). The peak at 112 amu corresponds to uracil$^+$.

The fragments are detected by a high efficiency multi channel plate (MCP) assembled with a secondary electron repeller grid in front of the first MCP. The time of flight (TOF) of all fragments created by one ion pulse is measured by a multistop time device and stored in an event-by-event mode. The TOF is derived from $T = T_0 \ p/(qE)$, where T_0 (time of flight of the fragment with no initial velocity) is proportional to M/Z, p is the projection of the momentum along the cell axis, and E is the extraction field.

In Fig.(5.7) (b) it is shown the time-of-flight spectrum of singly charged fragments and of two singly charged fragments measured in coincidence. Apart from the obvious difference in intensity, both spectra exhibit similar trends, with pronounced peaks for M/Z (in amu) 28, 42, and 69. The spectrum for non coincident singly charged fragments is very similar to that obtained with less energetic projectiles [203, 204]. This suggests that bond breakage is only efficient at specific locations within the molecule.

Comparison with theoretical and experimental conclusions have been done through two figures. The first one, Fig.(5.8), shows the agreement between the mass 2D ion/ion coincidence spectrum with theoretical mass coincidences of Table (5.1) given as junction of lines. The second figure, Fig.(5.9), shows the TOF 2D ion/ion coincidence spectrum where the signature of uracil^{2+} fragmentation has been zoomed (apart from H+ correlations, coincidences for M/Z $\leq$ 16 are much less intense). Combining the spectrum with the theoretical velocities for each fragment predicted on Table (5.1), we have generated a theoretical 2D TOF spectrum that is superimposed by color symbols. In order to obtain this spectrum, all the momenta of charged fragment are extracted directly from the theoretical values and are rotate in space randomly because isotropy (this is done many times in order to generate "all" the molecular orientations via Monte Carlo simulation). Only component along the axis cell is kept giving the a "spot" that, for the figure, it is limit to an open symbol.

As can be seen, the calculations predict the most intense coincidence signals, as well as (in most cases) their shape in the TOF distributions. This suggests that not only the fragments but also their relative velocities are correctly described by theory.

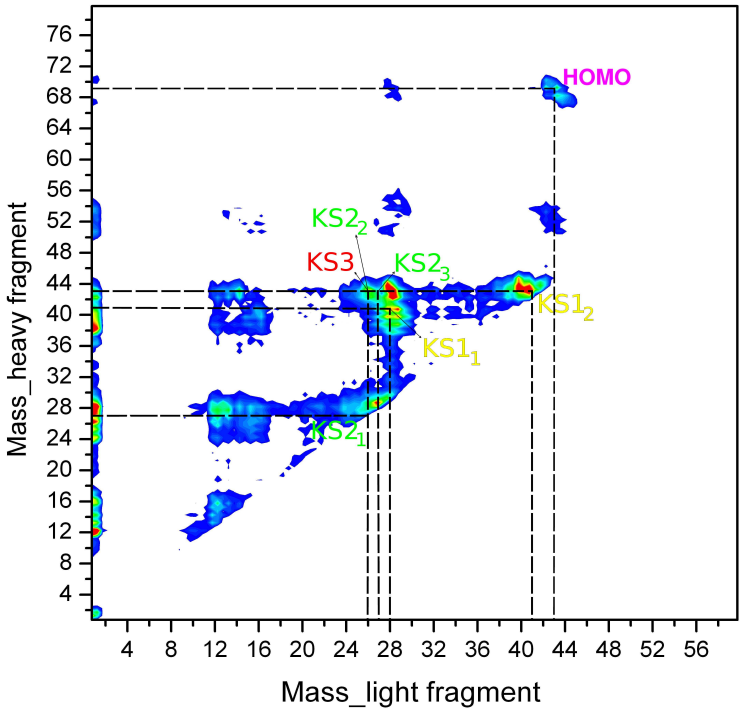


Figure 5.8: 2D coincidence mass spectrum of charged fragments after the proton-uracil collision. Results for KS1, KS2, KS3 and HOMO are pointed out by color letters.

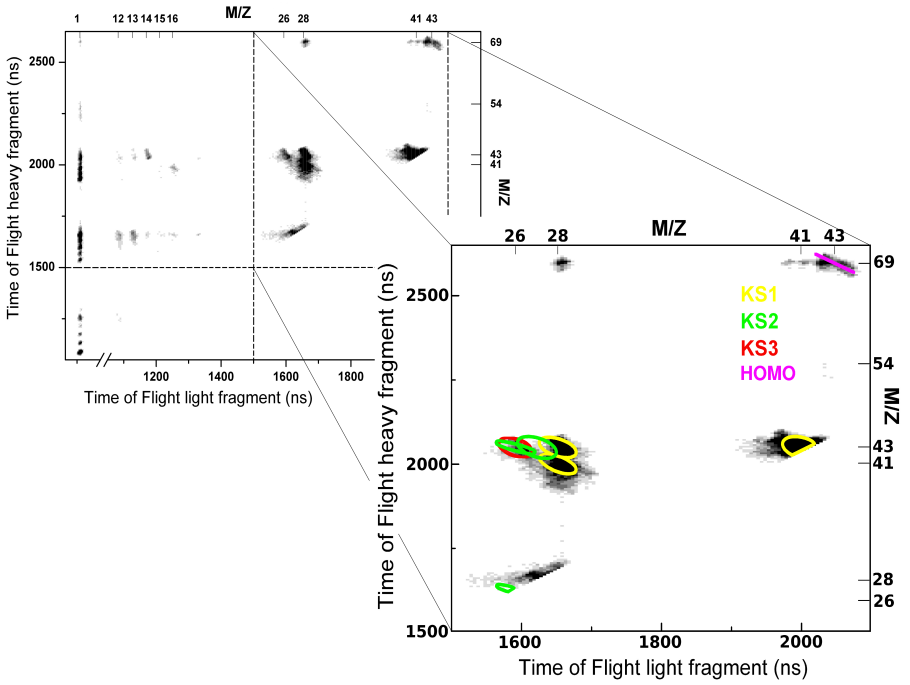


Figure 5.9: 2D TOF coincidence spectrum of charged uracil. In the inset, fragments for M/Z $\geq$ 22. Theoretical results for KS1, KS2, KS3 and HOMO fragmentations are represented by open symbols.

The energetically most favorable dissociation path, $HNCH_2 + CO^+ + OCNH^+$ (see appendix B.1), is only seen in HOMO fragmentation. Inner shell vacancies favor dissociation channels that lie much higher in energy but are dynamically favored by the special features of the electron density. Although these channels are associated with fission barriers of the order of 2-5 eV (see appendix B.2), the latter play a minor role due to the large amount of available energy ($\approx$ 50 eV). Therefore, one can conclude that the experimental observations are mainly due to sophisticated electronic dynamical effects taking place during the first few fs after removal of inner shell electrons.

In summary, a combination of TD-DFT MD and BO MD methods shows that two-electron removal from uracil leads to results in good agreement with ion/ion coincidence experiments in which gas-phase uracil is bombarded with 100 keV protons. The calculations show that fragmentation induced by vacancies in MOs of similar energy and/or localized in similar bonds strongly depends on chemical environment, which cannot be understood in terms of energetic stability of the fragments.

Chapter 6

Double ionization of Uracil in the liquid phase

Abstract *We apply a combination of Time-Dependent Density Functional Theory and Born-Oppenheimer molecular dynamics to study the early stages that follow the double ionization of uracil in liquid water. Ionization from equivalent Kohn-Sham orbitals than in gas phase lead to different fragmentation patterns. The solvent is revealed as an active participant in the dissociation process since it not only prevents the charge mobility of the biomolecule during the first femtoseconds of dynamics, but also drains positive charge from the uracil to outer hydratation shells. To evaluate the indirect effect of the radiation, double ionizations of a sourrounding water molecule of uracil are also considered. Ionizations either from one water molecule or the uracil lead to the formation of an atomic oxygen as a direct consequence of the molecule Coulomb explosion.*

6.1 Introduction

In order to evaluate the direct and indirect effect of the radiation, we set-up a system composed by one uracil molecule surrounded by 49 water molecules (total number of 159 atoms) in a cubic box of L=11.50 Å. A system of this size reproduces the density of liquid water at standard thermodynamics conditions, introducing two complete layers of solvent and part of the third hydration shell around the solute.

The exact number of molecules have been carefully considered following the preparation procedure of Ref. [205, 206]. Briefly, the cubic cell was first filled with water molecules only at ambient density of bulk liquid and the pressure was determined in a 500 ps run at a temperature of 300 K using standard classical force field MD. Afterwards, a cluster of H_2O molecules was replaced by the solute (uracil) and the pressure recomputed. Taking the pressure of the pure water sample as the target, the number of water molecules was adjusted until this pressure was recovered. The final configuration of this classical simulation was then used as the starting point for the TD-DFT Molecular Dynamics

Four cases are studied removing two electrons from different molecular orbitals, namely: two of them involve the removal of two different orbital localized on the uracil molecule (direct effect); in the other two, the ionization concerns orbitals of a neighboring water molecule situated 2.97 Å away from uracil (indirect effect). We will start off describing the two direct cases leaving the water ionizations to the end of this section.

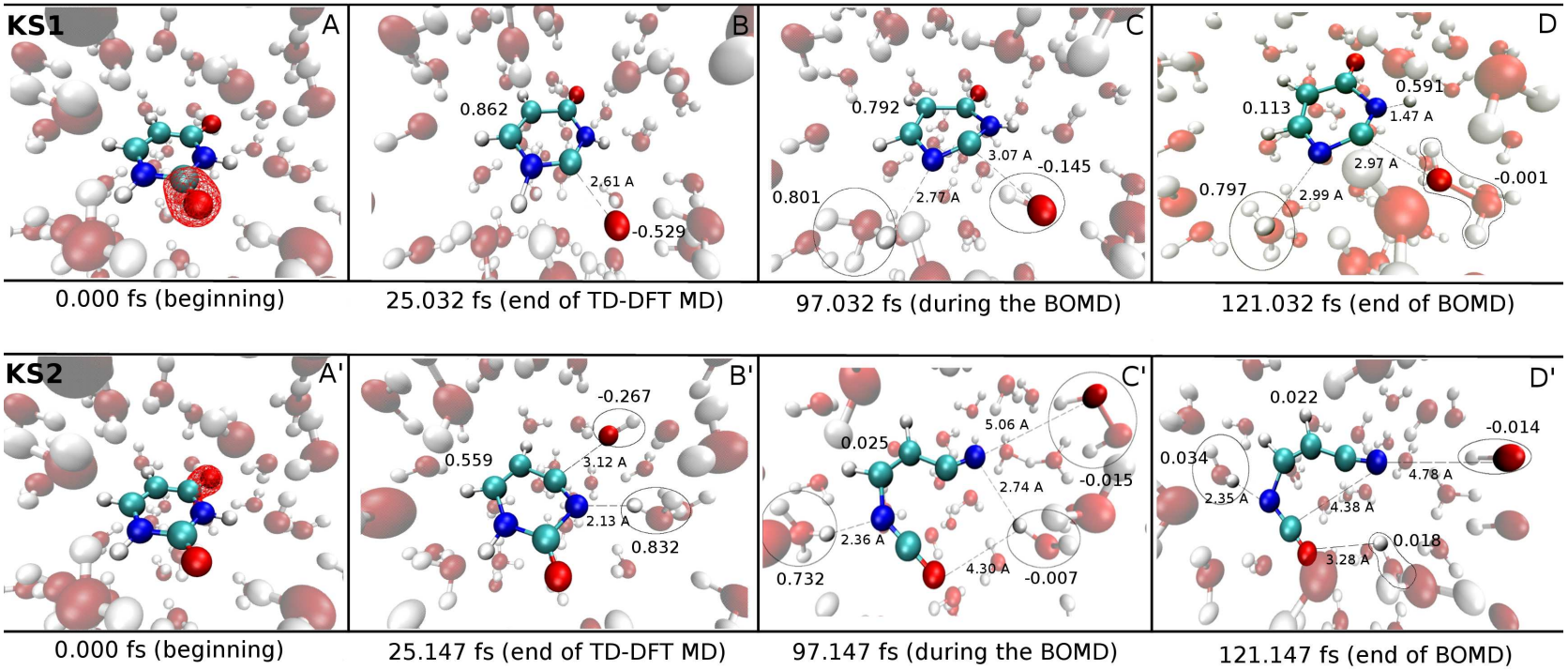


Figure 6.1: Snapshots corresponding to KS1 (upper strip) and KS2 (lower strip) dynamics in liquid phase. **A** panels correspond to the beginning of the simulation, representing the isodensity associated to each of the removed orbitals (isovalue 0.01). **B** panels are taken from the last step of the TD-DFT MD simulations showing in both cases a release of atomic oxygen. **C** and **D** panels represent and intermediate and final steps of the BO-MD. As it is showed, the ring generated after the TD-DFT MD simulation remains almost unaltered in the KS1 dynamics while it opens in the case of KS2. Some distances (in Angstron) and charges of the most representative species are given. For further details see the text.

For all simulations, the wavefunction is represented using periodic bounduary conditions and a cutoff of 70 Ry. Core electrons are replaced by pseudopotentials of the standard Troullier-Martins form [144], giving a total number of 436 electrons described explicitly. The Kleinman-Bylander [207] integration scheme is used for all atoms.

The computation in liquid phase results highly demanding. As a matter of fact, by including the solvent a TD-DFT propagation of 30 fs takes 36 Kh of CPU time, becoming 2.8 times more expensive than gas phase dynamics[1]. Nevertheless, *a priori* we can expect relaxation times to the ground-state to be shorter than in the gas phase since, in the liquid, excited uracil^{2+} is able to exchange potential energy by collisions and charge transfer with the surrounding molecules. Thus, time-dependent density propagation is reduced to 25 fs in the case of uracil ionizations, and to 10 fs for those involving water ionizations. For the later, the reduction have been done regarding the work Tavernelli et. al [34], where it is shown that double ionization of one water molecule in liquid water leads to the formation of one oxygen atom and two protons in less than 4 fs. The steps of integration are chosen to be 0.01 a.u. (0.00024 fs) for the direct effect and, it is increased to 0.05 a.u. (0.0012 fs) for the ionization of the water molecule.

As it is summarized in the appendix B.3, for indirect effect dynamics, a larger step does not affect the results at short times of propagation. All of these compromises have been carefully taken attending the better conservation of the energy at lower possible CPU time.

Last conformations at the end of time-dependent simulations have been followed by BO-MD dynamics (1 a.u. time-step), extended to a total time of 120 fs for uracil ionization dynamics, and to 58 fs for water ionization.

As we will see later, both types of dynamics lead to the release of a oxygen atom in the liquid. Theses species possess different physical properties that might produce different biological damages at major time scales. Comparing the results with the gas phase, uracil is less fragmented. At first sight, it can be consider as water plays a protective role, however, we must notice that it is in this medium where new ROS are generated and recombined.

6.2 Uracil ionizations

To clarify how the solvent affects the fragmentation patterns of uracil^{2+}, we must ionize the molecule in the same sites as we did in the gas phase. For this reason we looked in our system for two equivalent orbitals to KS1 and KS2 of the isolated uracil molecule that were centered, remember Fig. (5.1), in the two C-O bonds. The KS representation of our liquid phase gives a deepest orbital (-27.87 eV) similar to the one obtained in the gas phase. On the contrary, the second deepest KS orbital is not center in a C-O bond any more. Indeed, this orbital is partially delocalized between two water molecules making it useless for our purposes. Turning to Wannier representation 3.0.7 we find that orbital #20 (-26.56 eV) is centered over the C_4-O bond and is shape-equivalent to the KS2 in gas.For clarity, these two orbitals chosen will be referred as KS1 and KS2 according to their resemblance to gas phase uracil orbitals.

Main results are given by Fig.(6.1) in which snapshots of the upper and lower part show the KS1 and KS2 dynamics respectively. Panels **A** and **A'** represent the beginning of the dynamics with the ionized density plotted in red. **B** and **B'** panels show the end of the TD-DFT molecular dynamics simulation at 25 fs.

Thanks to the small step chosen, the non-conservation error of the total energy introduced by the numerical propagation represents only the 0.02% of the total initial energy. Following

[1]Values referred to 64 -IBM SP Power6 processors using CPMD version (3.13.2) compiled with Porland and IBM compliers with Lapack v3.2.1 and Essl v4.4 libraries.

the KS energy and the kinetic energy, the same behavior is found in both dynamics. As it can be seen in Fig.(6.2), there is a decreasing in the potential energy ≈ 8.5 eV during the first 5 fs. The temperature of the system reaches its maximum around ≈ 750 K at about this time. This similarity can be correlated with the release of an atomic oxygen that occur in both dynamics.

Finally, **C** and **D** panels of Fig.(6.1) display the situation after 72 fs and 96 fs of BO-MD concatenated from the end of the TD-DFT MD simulation. During the ground-state dynamics, some changes occur in the rings that remain after the oxygen loss.

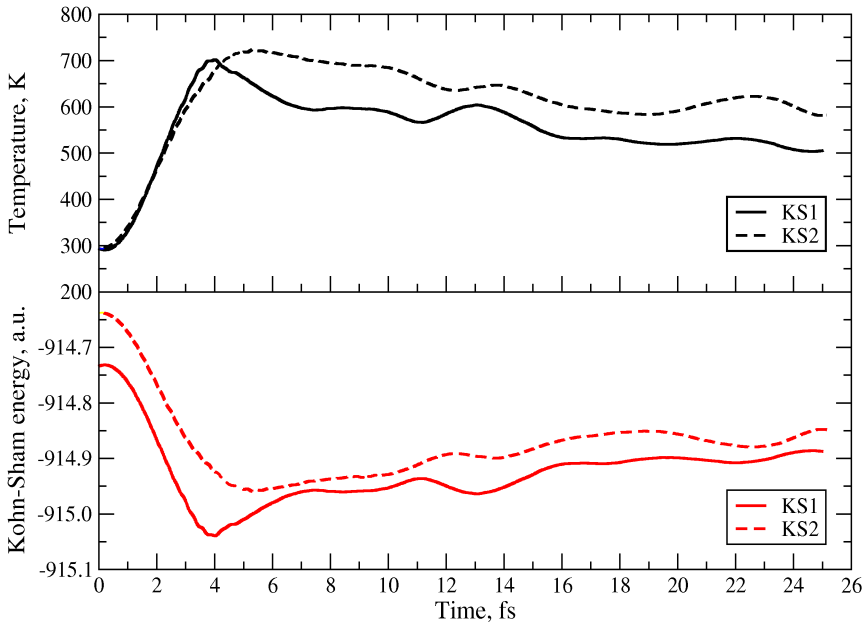
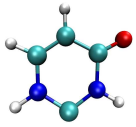
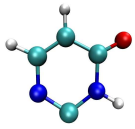
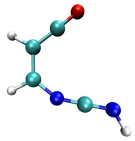
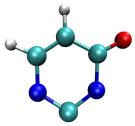
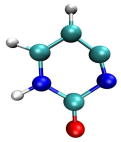
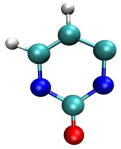
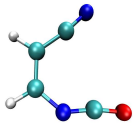


Figure 6.2: Time evolution of kinetic energy and Kohn-Sham energy of the system during the KS1 (solid lines) and KS2 dynamics (dashed lines)

To give more insights into these modifications several all-electron Gaussian09 [193] optimizations at BLYP/6-311+G(d,p) level have been performed over the intermediate and final structures. Results for gas phase and also using the Tomasi's Polarized Continuum Model (PCM) [208, 209] are summarized in Table (6.1). In this model, the solvent is considered as uniform dielectric medium (characterized by the dielectric constant ε_0) and the solute is place on its inside, within a cavity defined as the union of a series of interlocking atomic spheres. The polarization of the dielectric medium surrounding the solute is represented by a polarization charge density at the solute/solvent boundary. This charge polarizes the solute, and the solute and solvent polarizations are obtained self-consistently by numerical solution of the Poisson equation [210].

Although the method has been successfully applied to a widely variety of problems [211,212], the exclusion of a explicit solvent makes it fails for properties more sensitive to internal structure and dynamics of the medium (e. g. vibrational spectra [213]).

Table 6.1: Gaussian09 optimizations at BLYP/6-311+G(d,p) level, of the intermidiate and final structures of uracil appeared during the TD-DFT MD/BO-MD dynamics. Total energies of the optimized species in gas phase and using the PCM model phase are given in last columns. The third column indicates if the structure represents a minimum or not.

Name		Charge	Minimum?	E (a.u.)	E^{PCM} (a.u.)
KS1-Ring1		Neutral	YES	-339.5292286	-339.5450390
		1+	YES	-339.2370942	-339.3303684
KS1-Ring2		Neutral	YES	-338.9032326	-338.9129471
		1+	NO!	it opens to	*KS1-Ring2-open*
KS1-Ring2-open		1+	YES	-338.6527598	-338.7373960
KS1-Ring3		Neutral	YES	-388.2994210	-338.3070175
		1+	YES	-337.9819790	-338.0797323
KS2-Ring1		Neutral	YES	-338.9026939	-338.9187016
		1+	YES	-338.6246340	-338.7199912
KS2-Ring2		Neutral	NO!	it opens to	*KS2-Ring2-open*
		1+	NO!	it opens to	*KS2-Ring2-open*
KS2-Ring2 open		Neutral	YES	-338.3489367	-338.3580786
		1+	YES	-338.0021390	-338.0979813

KS1 dynamics

A more detailed view of most important steps occurring the KS1 dynamics is shown in Fig.(6.3). At 5.5 fs of TD-DFT MD simulation there is a bond breaking of C_2-O where the charge 2+ was initially located. In consequence an oxygen atom (O2) is violently shot out. The remaining ring is strongly rearranged to accommodate the charge and energy excess. There are observed, for instance, large N-H vibrational movements in the first fs of the dynamics ($\approx$ 18 fs). At about 25 fs, the oxygen atom is -0.529 charged and the rest of the uracil remains intact with a charge of +0.862. This species, that we name as *KS1-Ring1* in Table (6.1), represents a minimum for both gas phase and PCM.

During the ground-state dynamics, the H atoms (H1) attached to N_1 is transfered to a water molecule at 85 fs. The charge of structure that rests (*KS1-Ring2*) fluctuates between +0.5 and +0.8. Following the all-electron calculations, if the structure gets charged 1+, it does not represent a minimum and opens (*KS1-Ring2-open*). Thus, at 126 fs the ring looses the hydrogen atom (H3) that forms the N_3-H bond, leaving a finally six-membered structure (*KS1-Ring3*) with an almost neutral charge of +0.113. This structure results a minimum for gas phase and PCM optimizations.

In the meantime, the O2 released combines with an H atom of the medium forming an OH neutral species at 97 fs. The radical stays stable until is combined with another OH (previously formed) giving a neutral H_2O_2 at 126 fs.

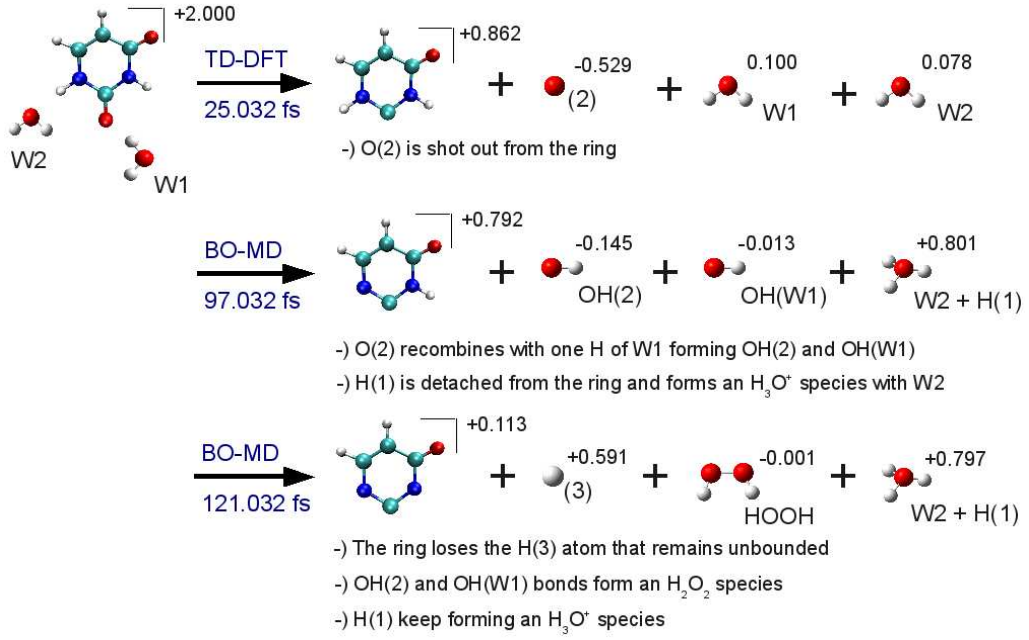


Figure 6.3: Summary of KS1 dynamics in terms of intermidiate and final species. The two most interacting molecules of solvent during the fragmentation are indicated and labeled by W1 and W2. These molecules are initially neutral.

Comparing the fragmentation with the one that occurs in the gas phase, the results have nothing to do. Isolated uracil gives three fragments of masses 43/41/28, while ionization in liquid leads to the formation of two protons (one forms an H_3O^+ species), one oxygen atom that combines to form different ROS's, and a structure that keeps a stable six-membered ring.

KS2 dynamics

As it is shown in Fig. (6.4) after the TD-DFT MD simulation the uracil not only looses the oxygen atom (O4) of the C_4-O bond where the initial charge was set; but also an H atom (H3) originally attached to the nearby N_3 atom. Both phenomena occur at almost the same time ≈ 5.2 fs. The uracil structure that remains (*KS2-Ring1*) represents a minimum for both gas phase and PCM, although it still changing due to dynamical effects. The resultant proton H3 is bonded to a neighboring water molecule and forms an H_3O^+ at 7.8 fs. Instead, O4 is linked to an H atom of a close water molecule in a longer process, ≈ 25 fs. The neutral OH species generated remains stable until it is combined to form a neutral H_2O_2 at 77.7 fs. The new peroxide decomposes at ≈ 100 fs leading to the former OH.

As it happend in the previous case of KS1, during the BO-MD there are some changes over the structure that stays behind after the oxygen release. At 37 fs, H1 is detached from N_1. Soon after, ≈ 41 fs, there is a bond breaking between C_2 and N_3. The opening of the ring is in agreement with all-electron calculations, that is, after the second proton loss the neutral structure generated (*KS2-Ring2*) is not a minimum, neither gas phase nor PCM calculations, and get opened (*KS2-Ring2-open*). Both calculations predict a final structure in which the breaking is done in the same site as it happends during the dynamics.

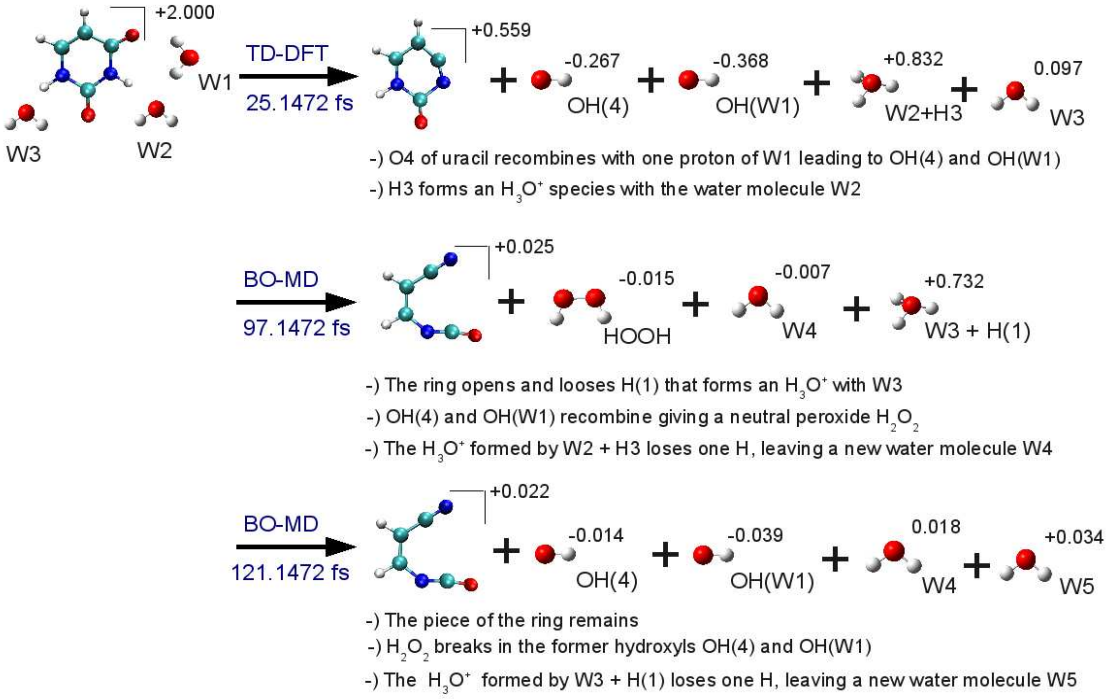


Figure 6.4: Summary of KS2 dynamics in terms of intermidiate and final species. Most interacting molecules of solvent during the fragmentation are indicated and labeled by W1, W2 and W3. These molecules are initially neutral.

To conclude, double ionization of KS2 in liquid water produces one OH, two protons and a neutral opened structure that maintains the former six atoms of uracil's ring. This conclusion differ from the one obtained in gas phase where four fragments of masses 43/27/26/16 were generated.

6.3 Water ionizations

If a cell is exposed, the probability of the radiation to interact with the DNA molecule is very small because these critical components make up a very tiny part of the cell. Certainly, the highest probability is given by the interaction with water since each cell is mostly composed of it (about the 80% in mass). When radiation affects the water, it may break the bonds that hold the molecule together, producing fragments such as hydrogen and oxygen species. These products may recombine or may interact with other fragments or ions to form compounds which would harm the cell.

The experimental and theoretical study of water radiolysis in swift ion tracks has received a lot of attention in the recent years [29]. It provides remarkable information about how the secondary species generated after the ionization could combine to form toxic substances. Unfortunately, the ionization by itself and the subsequent chemical processes that occur at the femtosecond time-scale are experimentally unaccessible so far. The only way to truly approach these first events is by using theoretical methods based on time-dependent molecular dynamics.

As it has been referred previously, first investigations of our group using TD-DFT MD simulations concerned the double ionizations of a selected water molecule in sample liquid water [34]. This time however, our main concern is frameworked in the so-called quasi-direct effect [3,4,214], that is, we are interested in evaluate the effects of such ionization when it takes places in a surrounding water molecule of the first hydratation shells of uracil. Our concern will be to check if the immediately protons caused by the coulombian explosion induce any variation on the uracil's structure, and also to study the physical properties of oxygen species generated.

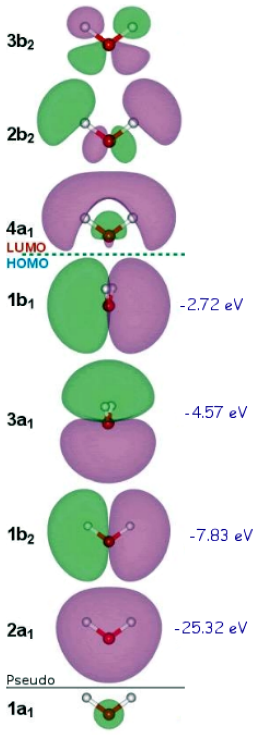


Figure 6.5: Molecular orbitals of water

With this purpose on mind, we run two dynamics ionizing a water molecule situated 2.97 Å to the nearest uracil's atom (specially the oxygen atom of the C_2-O bond). This molecule belongs to the second hydratation shell and it is not bounded to the biomolecule.

The ionizations are done by removing two electrons from a different molecular orbital of the chosen water molecule each time. As it is shown in Fig.(6.5) each water molecule have five occupied molecular orbitals referred following their symmetry properties as: 1a1, 2a1, 1b2, 3a1 and 1b1. In our pseudo-potential representation the core orbital 1a1 is not simulated, leaving only four orbitals (eight electrons) treated explicitly. For their representation, Wannier Orbitals centered on each O atom are used. This type of orbitals enable to separate molecular orbitals of each water molecule from the ones of the biomolecule.

For the water molecule chosen, the corresponding molecular orbital energies are: -25.32 eV, -7.34 eV, -4.57 eV and -2.72 eV respectively. It is from the two deepest orbitals, i.e. 2a1 and 1b2, where the ionizations are proposed. The gap of $\approx$ 18 eV existing between them will allow us to estimate the influence of the energy deposited in the initial H_2O^{2+} ions.

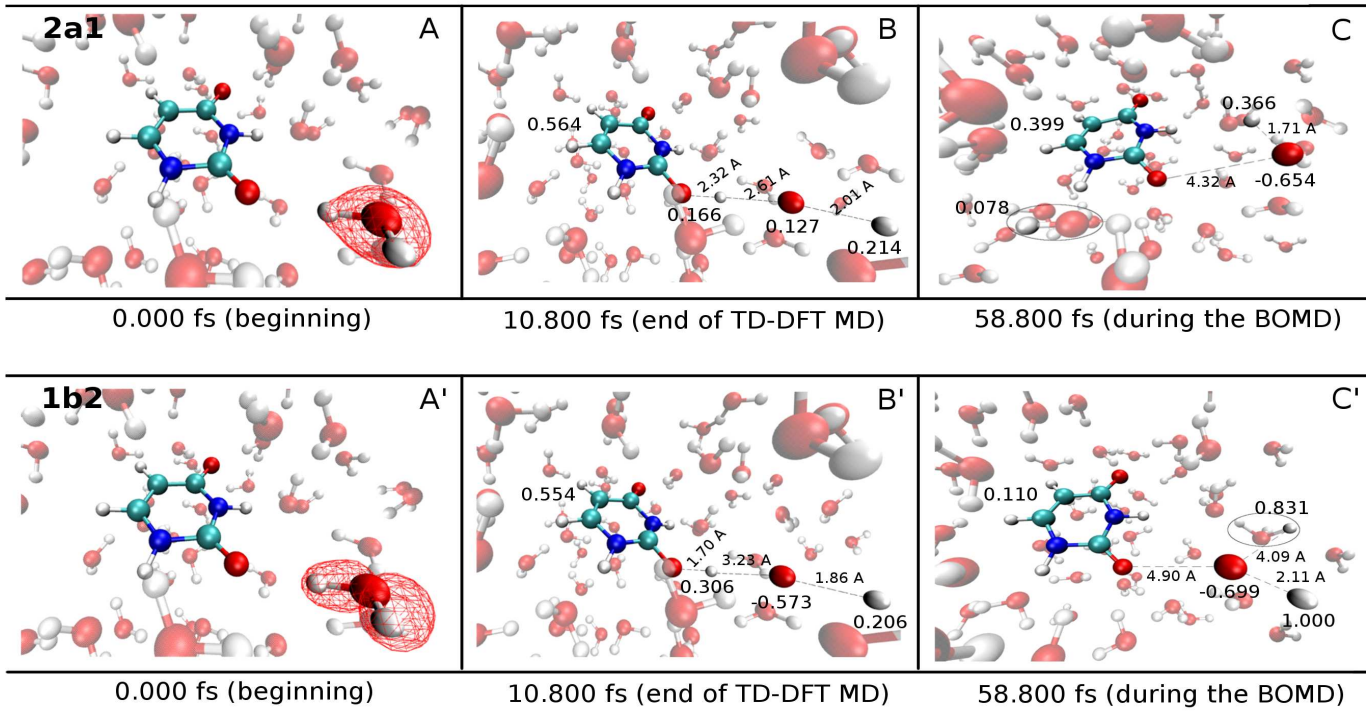
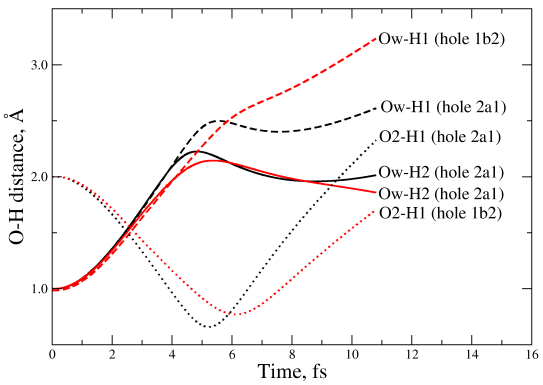


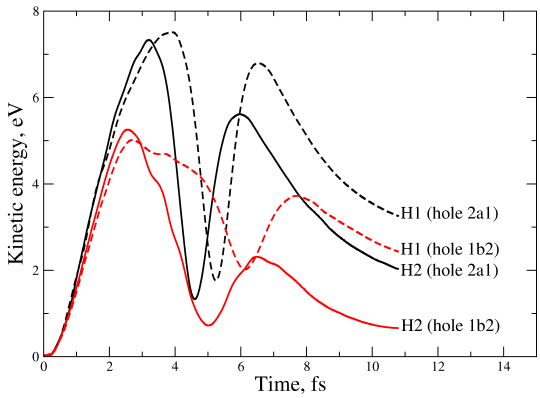
Figure 6.6: Snapshots corresponding to 2a1 (upper strip) and 1b2 (lower strip) water double ionizations. **A** panels refer to the beginning of the simulation, representing the isodensity (isovalue 0.01) associated to each of the removed Wannier orbitals. **B** panels are taken from the last step of the TD-DFT MD simulations, and **C** panels are taken from the final step of the BO-MD. For all, some distances (in Angstron) and charges of the most representative species are given.

A summary of the results is given by Fig.(6.6). After 10.8 fs of TD-DFT MD both cases represent almost the same situation: there is a coulombian burst of the H_2O^{2+} that leads to the formation of one oxygen atom and two protons (panel **B** and **B'** on the figure). No big changes occur in the uracil that remains unaltered and slightly charged (a half of a charge in both cases). After the ground-state simulations (**C** panels) the ring charges are reduced. The released oxygen atoms finally get -0.654 and -0.605 charged. The characteristics of these atoms will be discussed in next section.

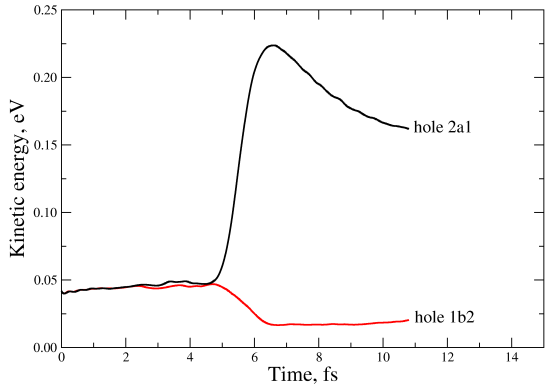
Concerning the ejected protons, most relevant information during the TD-DFT MD simulation is depicted in Fig.(6.7). In both simulations there is proton that collides the uracil (from now on H1) and the other one (labeled as H2) that interacts with a neighboring water molecule. As it is showed in panel (a), the O-H bond breaking is produced in similar ways at the same time ($\approx$ 3 fs) for the two ionizations. Small differences are found, however, in the trajectories. For both dynamics the proton H1 that goes towards the uracil covers a bigger distance than H2 proton. The later is slowed down by interactions with neighboring water molecules. The protons that target the biomolecule collide with it and hit the oxygen atom of the C_2-O group. No attachment are detected, the atoms collide and bounce in a process that lasts 7.5 fs for the case of 2a1, and a bit longer, 9.0 fs, for 1b2 ionization.



(a) O-H distances



(b) Kinetic energies of the ejected protons



(c) Kinetic of the mass center uracil molecule

Figure 6.7: Data obtained from the TD-DFT MD simulation of 2a1 (black lines) and 1b2 (red lines) water ionizations. Panel (a) shows by solid and dashed lines the O-H distances between the oxygen (Ow) and the two protons of the ionized water molecule. Dotted lines represent the distances between the O2 of the uracil and the colliding protons. Panel (b) displays the kinetic energies of the released protons and panel (c) the kinetic energies of the mass center of the uracil.

Panel (b) attends to the kinetic energies of the ejected protons during the dynamics. The released protons coming from the 2a1 water ionization present higher values than the ones liberated in the 1b2 dynamics. As it is reflected on the figure, protons shot out from the deepest water molecular orbital get maximum kinetic peaks of $\approx$ 7.5 eV. On the contrary, the second deepest orbital protons reach maximum values of $\approx$ 5 eV. These differences are related with the initial energy of the double ion at the beginning of the dynamics. In all cases, sudden variations

are present mainly due to the interactions with water and uracil molecules. Nevertheless, the major difference is found in the red dashed curve since the proton H1 coming from 1b2 hole interacts with the uracil in a diverse way than 2a1. In fact, the collision angles described by the entry/exit trajectories are much different: 42° and 83° for 2a1 and 1b2 respectively. It turns out there is a more scattered trajectory for 1b2 than 2a1 in which the collision is more direct. For instance, the proton ejected in the 2a1 dynamics reaches a closest distance to uracil equal to 0.66 Å, at 5.2 fs, while less energetic case 2b1 gets a minor distance of 0.77 Å at 6.1 fs.

Fig.(6.7) gathers the kinetic energy of the mass center of the uracil during the time-dependent dynamics. As it can be seen, the energy loss of the protons is not transmitted as kinetic energy to the uracil. Actually, only the deepest ionization (black line) transfers kinetic energy to the biomolecule ($\approx$ 0.2 eV). The collision is far from being elastic, aiming to almost entire potential energy exchange (charge transfer for instance) between the protons and the biomolecule.

We can conclude that the ionization of a surrounding water molecule of uracil steers to the formation of a negative oxygen atom and two protons. The energy of the ejected protons are between 5 and 7.5 eV in accordance with the energy of the ionized orbital. Spite of some of them collide with the uracil, no significant changes are produced in the structure of the biomolecule. The possible biological damage of such ionization is then reduced to the indirect effect that the released oxygen atoms might cause.

6.4 Liquid vs gas phase

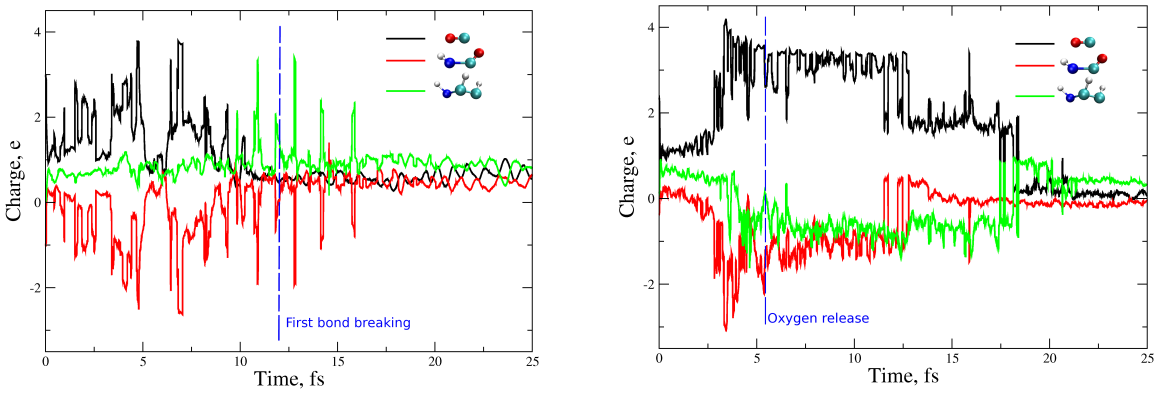
As we have already referred, different fragmentation patters are found in the liquid than in the gas phase. Ionizations from the two deepest KS1 and KS2 orbitals of the isolated system cause the fragmention into three an four fragments. The ion embedded in water, on the contrary, undergoes softer fragmentations that help to keep bigger fragmenta (for instance KS1 ionization preserves the six-membered ring of the uracil, Fig.6.3)

These dissimilarities are closely related with the role of solvent, specifically with the charge transfers that occur with the solute. The water-uracil interactions not only helps to localize the charge within the biomolecule, but also they help to migrate the positive charge of the uracil to outer hydratation shells.

In this sense, Fig.(6.8) and (6.9) recap the information attending to atomic charges as direct measuring of the density evolution. The first figure is divided into two panels (a)-gas and (b)-liquid representing the charge of the uracil during KS1 TD-DFT MD for both phases. To make it simpler, the uracil has been divided into three fragments that corresponds to the ones obtained in the gas phase dynamics: 28/43/41 or CO, HNCO and HNCHCH. Although graph (b) represents fake fragments, they are useful for our purposes.

Looking at the figure, both phases represent different charge distributions because, of course, in both the way of fragmentation is different. Indeed, the first bond breaking in gas takes longer (N_1-C_2 at $\approx$ 14 fs, see Fig.5.3) than in the liquid, where at 5.5 fs an O atom is ejected. To compare the charge mobility that occur into the biomolecule we should regard only the time where molecule does not fragment, that is, times below 5.5 fs.

For both and just at the beginning, the initial charge allocated in the CO is quickly reduced by charge exchange with the other two fragments. Afterwards, big fluctuations occur in gas ionization: the charge of the CO is suddenly and directly exchanged with HNCO. On the other hand, the HNCHCH fragment remains almost neutral. For liquid phase, the charge over the CO keeps localized around +1 during the first femtoseconds without sudden changes. The effect is attributed to the presence of water since it is the only element that differs from one dynamics to



(a) KS1-gas phase dynamics

(b) KS1-liquid phase dynamics

Figure 6.8: Comparison of the uracil charge in terms of fragments for gas (a) and liquid (b) TD-DFT MD. The charge over the ionized C-O piece is given by the black line. The rest of the ring is separated into two fragments: HCNO, red line, and HNCHCH, green line.

the other. After 3 fs, when the oxygen atom starts to detach, the charge is gradually increased as the ring drains charge from the C atom.

As we have already told, the influence of the solvent is not only reflected in the charge mobility inside the uracil molecule, but also it plays role as a drain of positive charge. To measure this effect we have defined the hydratation shells of our system attending to the distances that separate the mass center of the uracil and the O atoms of the water molecules. If the distance R is ≤ 4 Å the molecule is included in the first shell (3 molecules), if 4 Å $<$ R ≤ 5 Å (11 molecules), they are in second shell. Finally, molecules with 5 Å $<$ R (35 molecules) are counted together (Ask MP if the separation is ok)

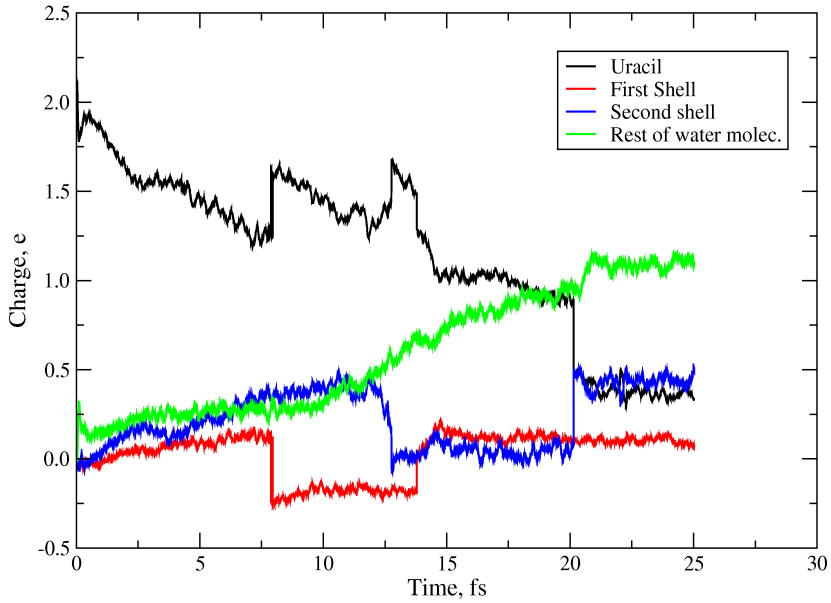


Figure 6.9: Charges of uracil and different hydratation shells during the TD-DFT MD.

Thus, black line of Fig.(6.9) represents the charge of the atoms of the uracil as time evolves. There is a descent trajectory with four abrupt changes produced when the protons are detached/attached by the reverse movement after the oxygen ejection. At the end of the dynamics,

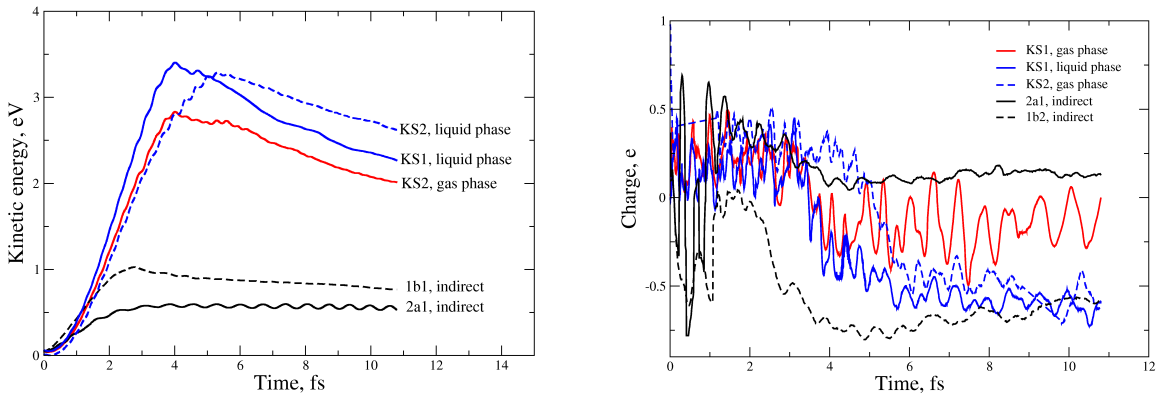
the molecule gets +0.3 charged as the sum of an negative oxygen atom (-0.5) and a positive piece (+0.8).

In the meantime, while the uracil is getting neutral the charge rises in the solvent. Specially, the outer shells of hydratations are the ones that become more positively charged. As it is shown in the figure: the first hydratation shell (red line) barely changes its initial charge, the second shell (blue) gets a half o a charge and the outermost (green) is +1 charged. It is important to emphasize that during the dynamics there is no H^+ production by the uracil, the charge is transferred only by electrostatic interactions with solvent.

About the similarities between the two phases, we have found that there is a oxygen release in both. Namely, the ionization of the KS2 in gas and ionizations in the liquid (all direct and indirect cases) lead to the formation of one atomic oxygen. Attending to their physical properties, the meaning of Fig.(6.10) is to show how different these species are.

Thus, left panel (a) shows the kinetic energies. It could be said that the oxygen arisen from the ionization of a water molecule (black lines) represent a much less energetic profile. In fact, the kinetic energy peaks are below 1 ev. Notice that for the case of the deepest ionization 2a1, the energy is even less than 1b2. As it was depicted in the second graphic of Fig.(6.7), this fact is due to a major energy transference to protons.

Surprisingly, looking at the oxygen charges in panel (b) of Fig.(6.10), it can be seen that both oxygen are not the same: the one released by the 2a1 ionization gets an almost neutral charge while the one coming from the 1b2 ionization seems to be a negative radical.



(a) Kinetic energies

(b) Charges

Figure 6.10: Kinetic energy (a) and charges (b) of different oxygen atoms released during the TD-DFT MD

The oxygens ejected by uracil ionizations in gas (red line) and liquid (blue lines) also have characteristics in common. For instance, their maximum kinetic energies are in between 2.5-3.5 eV at 4-6 fs. The two cases in liquid, KS1 and KS2, lead to negative oxygens whereas the KS2-gas oxygen also gets negative charged, but in this case, it varies between 0 and -0.25. This particular difference is due to the absence of solvent. In the liquid once the oxygen is released, it starts to drain negative charge from the nearby water molecules.

Chapter 7

Double ionization of small water clusters

Abstract *We study the dissociative ionization of water clusters by impact of 12 MeV/u Ni^{25+} ions. Cold target recoil ion momentum spectroscopy (COLTRIMS) is used to obtain information about stability, energetics and charge mobility of the ionized water clusters. An unusual stability of the $H_9O_4^+$ ion is observed, which could be the signature of the so-called Eigen structure in gas-phase water clusters. From the analysis of coincidences between charged fragments, we conclude that charge mobility is very high and is responsible for the formation of protonated water clusters, $(H_2O)_nH^+$, that dominate the mass spectrum. These results are supported by CarParrinello molecular dynamics and Time-dependent Density Functional Theory Simulations, which also reveal the mechanisms of such mobility.*

7.1 Introduction

The past decade has brought tremendous advances in the study of molecular clusters. One of the goals has been to understand how the properties of these molecular clusters vary with size and, in particular, how condensed phase attributes develop as the size increases [215–219]. Most investigations have focused on either weakly bonded van der Waals systems, like rare gases, or strongly covalently or metallically bonded systems. Water clusters belong to a different category in which the individual molecules interact through hydrogen bonds whose strength lies between that of weak van der Waals and strong covalent clusters [220].

The dynamics of water clusters is interesting for a number of reasons. In atmospheric chemistry, it is important tounderstand the early stages of cloud and droplet formation [221]. Water clusters have also been suggested as transient intermediates in liquid water [222], which is crucial to understand many chemical and physical processes [220]. In particular, ionization of water clusters is relevant to radiation damage studies because water is the natural environment of biomolecules. It has been shown, e.g., that such an environment may efficiently protect the biomolecules against ionizing radiations [199] or, in contrast, lead to secondary electrons and radicals (water radiolysis) that react with the biomolecules.

In this work, we study the fragmentation dynamics of small water clusters in collision with highly ionizing 12 MeV/u Ni^{25+} ions by using imaging techniques with coincident measurement of the full momentum vector of the charged fragments. This technique gives us simultaneous information on stability, energetics and charge mobility inside the cluster. The measurements

are complemented by CarParrinello molecular dynamics (CP MD) and Time-Dependent Density Functional Theory (TD-DFT) simulations that allow us to unambiguously conrm the experimental ndings.

Water clusters are produced using a supersonic expansion. At the start of the experiment, approximately 20 ml of distilled water is loaded into a heated reservoir (stagnation chamber). Its temperature is controlled by a thermocouple and is typically kept at $T_S \approx 80°C$. The pressure of the water vapour p_s is determined by the temperature of the stagnation chamber, and is around 400 mbar. A 1 bar Ar seeding gas is mixed with the water vapour at the exit of the sealed oven to increase the total stagnation pressure. The vapour ows to the nozzle (30 μm diameter, temperature 90°C) through a stainless steel tube. The temperature of this tube is controlled separately, and is held slightly above the reservoir temperature to avoid condensation. The gas mixture undergoes isentropic expansion and the temperature drops rapidly with an increasing distance from the nozzle. This leads to the supersaturation of the water vapour and subsequent clustering [223]. Note that clusters produced by such an adiabatic expansion are usually considered to be close to the cluster melting temperature [224], but the seeding with argon may result in colder clusters. Under the present conditions, small size clusters are expected to be produced [225]. 12 MeV/u Ni^{25+} projectiles produced by the GANIL facility in Caen intersect the $(H_2O)_n$ supersonic cluster beam. The same projectile has been used in our previous study of dissociative ionization of a single water molecule [226–229]. The charged fragments are extracted by a uniform electric eld with 4π solid angle acceptance and directed onto an 80 mm diameter position sensitive detector. As in the case of molecules [227, 230–232], the energy distribution of the fragments is determined from both time of ight and coordinates of the impact position of each particle detected. Such coincidence experiments are necessary to unravel the repartition of the initial multiple ionization on the subsequent fragments.

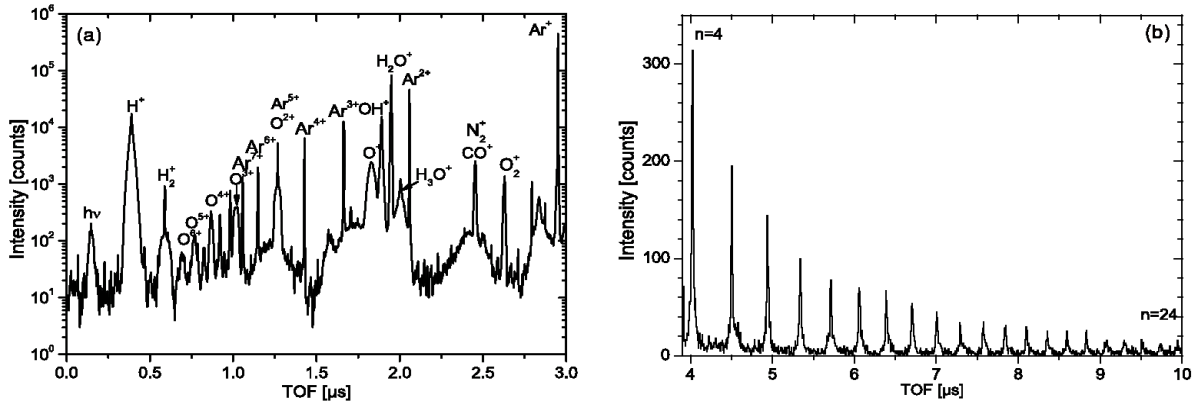


Figure 7.1: Time-of-ight spectrum for 12 MeV/u Ni^{25+} on $(H_2O)_n$ (a) For short time of ight (less than 3 μs) and (b) for long time-of-ight (more than 4μs) where protonated $(H_2O)_nH^+$ are detected.

An important issue is to know whether the clusters are stable at the temperature they are produced in the experiment. For this purpose, we have carried out CP MD simulations of clusters containing 3 and 11 water molecules at different temperatures. All simulations were performed with the plane wave KohnSham (KS)-based DFT code CPMD [100, 152]. We have followed exactly the same procedure as in [31, 34] for the study of liquid water. However, in the present work, the water clusters have been enclosed in a much larger cubic box of length L = 21 $\AA$ to avoid interaction between clusters in neighbouring boxes. The energy cut-off of the plane wave basis is also larger: 90 Ry. The optimized initial cluster geometry for $(H_2O)_3$ is in good

agreement with that reported in [233]. The results of these calculations show that, after 0.7 ps, $(H_2O)_3$ is stable below 90 K while $(H_2O)_{11}$ is stable below 300 K. Since clusters produced in the supersonic expansion are rather cold (T < 80 K) [234], we infer from thermal stability of the neutral clusters that the ionization process occurs before fragmentation processes. Some small clusters are therefore present in the jet, which could be further ionized and their product detected. Time-of-ight spectra are presented in Fig.(7.1). They allow us to identify the species produced in the collision. For short time of ight (less than 3 μ, Fig.(7.1)(a)), the spectrum is very similar to that obtained for a single water molecule. This is not surprising because, under the present experimental conditions dominated by the formation of small clusters, individual water molecules are also produced in the supersonic expansion. We detect multiply charged oxygen fragment O^{q+} ions (with q ranging from 1 to 6) coming from dissociation of H_2O^{8+} molecular ions produced in a single collision event (O^{6+} ions are always emitted in coincidence with two protons). We also detect Ar^{q+} ions coming from the seeding gas ionization.

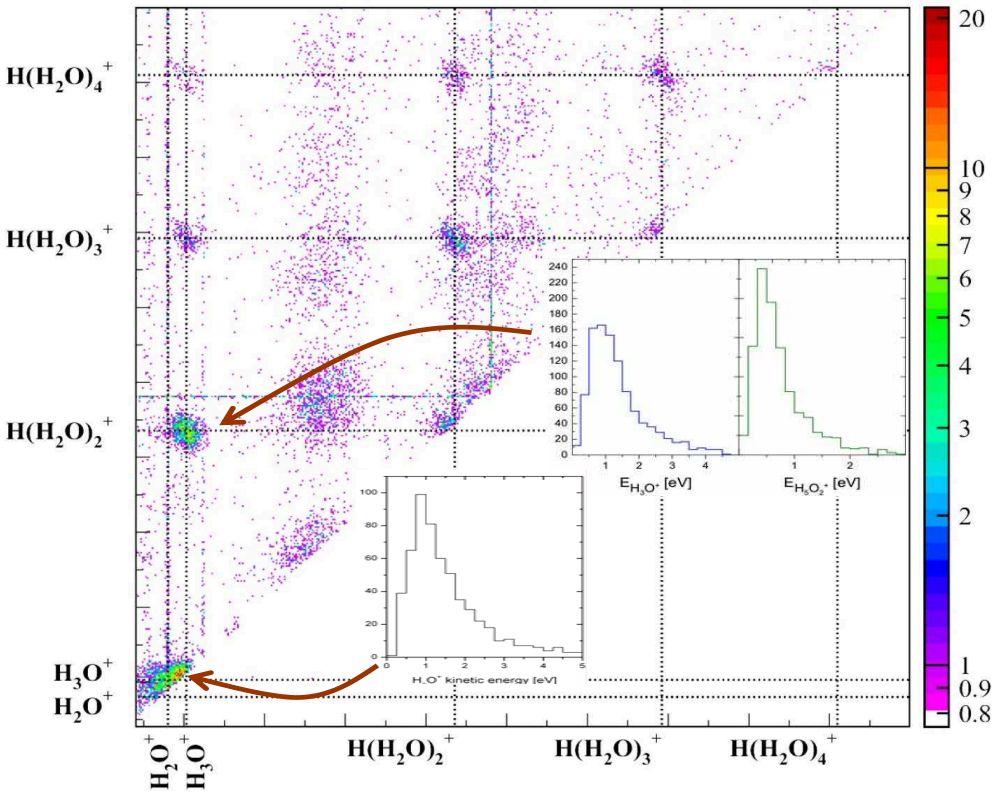


Figure 7.2: Coincidence spectrum between singly charged fragments. The insets show the kinetic energy distribution of coincident H_3O^+ and $H_5O_2^+$, and of two coincident H_3O^+ cluster fragments for 12 MeV/u Ni^{25+} + $(H_2O)_n$ collisions.

Cluster ions appear clearly for longer time of ight (more than 4 μs, Fig.(7.1)(b)). In this case, the mass spectrum is dominated by protonated cluster ions $(H_2O)_nH^+$ (or $(H_2O)_{n-1}H_3O^+$) and reects, in some way, the initial size distribution produced in the supersonic expansion. Indeed, the formation of $(H_2O)_nH^+$ is dominated by large impact parameters, which produce mainly singly ionized clusters. However, it is difcult to establish a direct link between this spectrum and the initial size distribution. To check whether H_3O^+ fragments can also arise from singly charged $(H_2O)_n^+$ water clusters produced in the collision, we have performed CP MD calculations for $(H_2O)_3^+$ and $(H_2O)_{11}^+$ during 0.7 ps for several temperatures. We have found that the larger

cluster is stable up to T = 200 K, while the smaller cluster breaks even at T = 10 K. When the latter breaks, a smaller $(H_2O)_n^+$ fragment is produced but never $(H_3O)+$. This suggests that a minimal size is required to observe such a fragment. We do not detect doubly charged cluster fragments. Previous photoionization experiments have determined that a critical cluster size of n = 37 has to be reached to observe series of the form $(H_2O)_nH^{2+}$. This critical size is smaller in collisions with highly charged ions [235]. Such a difference is generally attributed to the fact that, in a collision, double electron removal takes place at a large impact parameter, thus leading to "cold" dications that can survive more easily. In the present experiment, the critical cluster size is clearly above n = 20.

In Fig.(7.3), we have plotted the relative intensities of the peaks shown in Fig.(7.1)(b) as functions of size n. To a good approximation, the intensities follow an exponential behaviour, with the notable exception of $(H_2O)_4H^+$. Similar exponential behaviour has been observed in cluster growth studies [225]. The unusually high stability of $(H_2O)_4H^+$ has also been reported in the case of free jet expansion of liquid water [236] and for clusters sputtered from frozen H_2O under keV He^+ [237] and MeV/u Ar ion impact [238].

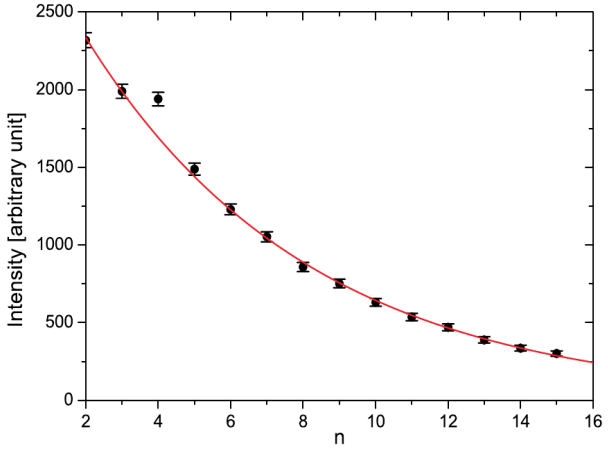


Figure 7.3: Relative intensity of the different protonated $(H_2O)_nH^+$ species as a function of n. The line is an exponential data fit.

To our knowledge, this is the first time that such enhanced stability is seen from gas-phase water clusters that are subsequently ionized. $(H_2O)_4H^+$ clusters have been extensively studied [239] since they are possible building blocks for large proton hydrate species. There is continued speculation about the location of the excess proton in these hydrates and, by extension, in the bulk [240]. Most chemists consider that the hydrated proton appears in the form of a hydronium ion, H_3O^+. But this picture was rened by Eigen [241] and Zundel and Metzger [242], who advocated larger complexes such as $H_9O_4^+$ and H_5O^+, respectively. In the "Eigen cation", the central hydronium ion is strongly hydrogen-bonded to three H_2O molecules forming $H_3O^+(H_2O)_3$. In the "Zundel ion", the proton lies midway between two water molecules H_2O – H^+ – H_2O. The most appropriate picture to describe the hydrated proton is still a subject of debate. The relatively high stability of the $(H_2O)_4H^+$ fragments observed in this work seems to support the Eigen cation as the preferential conguration.

In order to explore the charge repartition on the cluster fragments, we have analysed in detail the coincidences between the different charged species. A relevant region of the coincidence spectrum is shown in Fig.(7.2). We unambiguously see coincident emission of singly charged protonated cluster fragments. In general, the dominant coincidence events involve at least one H_3O^+. High resolution energy and angular distributions of the charged fragments are determined

from the measured full momentum vectors. Energy distribution measurements are still scarce in the case of protonated water clusters [234, 243]. However, precise measurements are necessary as dissociation energies can vary depending on the dissociation path [244]. As an illustration, the insets in Fig.(7.2) show energy distributions for H_3O^+ and $H_3O^+/H_5O_2^+$ coincident ions and for two H_3O^+/H_3O^+ coincident ions (in the latter case, a single distribution is shown because both are almost identical). From the analysis of the mass spectra for a given number of charged particles per event, we can conclude that the two coincident ions mainly come from the dissociation of a doubly charged parent in two (and only two) singly charged fragments plus eventually some neutrals. The H_3O^+ and $H_5O_2^+$ coincident distributions peak around 900 meV and 500 meV, respectively, while the H_3O^+ and H_3O^+ ones peak around 900 meV. In the latter case, our measured angular distributions show that the two H_3O^+ fragments are preferentially ejected at 110°-120°C , thus evidencing that another neutral fragment(s) contributes signicantly to the total momentum balance.

These findings lead us to explore the question of chargelocalization and mobility within the multiply charged water clusters. In this respect, we never detect multiply charged oxygen ions (even O^{2+}) in coincidence with cluster fragments, even after five-fold ionization of the cluster. This suggests a high charge mobility similar to that observed in multiply ionized van der Waals clusters of fullerenes [245]. It also agrees with synchrotron radiation ionization experiments in which resulting holes are rather delocalized with a preference for molecules lying in the cluster surface [39].

7.2 TD-DFT MD simulation of $(H_2O)_3^{2+}$ and $(H_2O)_{11}^{2+}$

To elucidate the mechanisms behind this charge mobility, we have performed TD-DFT MD calculations of the fragmentation of the doubly charged clusters $(H_2O)_3^{2+}$ and $(H_2O)_{11}^{2+}$ produced by removal of two electrons from the corresponding neutral clusters. We have chosen the 11 water cluster because it is the smallest size with a cubic closed cage and an attached cycle. Such a conguration allows three different coordination numbers depending on the position of the water molecule in the cluster. This has allowed us to study the effect of ionizing a tri-coordinated water molecule (A in Fig.(7.4)) or a tetra-coordinated one (B in Fig.(7.4)). In these TD-DFT MD calculations, we have used the same procedure as in [31, 34] for electron removal. We have also evaluated the Bader charge of the different species produced in the fragmentation process. For the larger cluster, the main results of the simulations are presented in Fig.(7.4). The left panel shows the initial geometry of the $(H_2O)_{11}^{2+}$ cluster, which is the same as that of the corresponding neutral cluster for a Frank-Condon transition. Two cases have been considered: both electrons are removed from either the A or the B molecule (see the left panel). In both cases, two electrons were removed from the 1b1 orbital which is one of the most likely processes. It is worth noting that direct two-electron removal or the Auger process will lead to very similar two-valence hole conguration. The middle and right panels in Fig.(7.4) show the corresponding cluster geometries after 12 fs.

In case A, two H_3O^+ ions are clearly produced (in both cases, the calculated charge $> +0.8$). These ions lead to one of the most intense peaks in Fig.(7.2). The H_3O^+ ions do not arise from the water molecule that loses the two electrons, but from water molecules that "capture" protons ejected from the ionized molecule. The capture proceeds through several steps that involve neighbouring water molecules in which the protons are successively captured and released. This is very similar to what has been found for liquid water [31,34]. The dynamics is very similar when electrons are removed from molecule B. In this case, only one H_3O^+ ion is produced (calculated charge $> +0.8$), but again it does not come from the water molecule that

was originally ionized. Our simulations for the smaller doubly charged cluster $(H_2O)_3^{2+}$ do not lead to H_3O^+ fragments, thus implying that there is a critical size for the production of these fragments. These results are at variance with those found in [246], where highly charged oxygen ions O^{q+} ($q \leq 6$) were abundantly produced upon laser-field ionization of large water clusters (hundreds of molecules). This means that, in this case, Coulomb or hydrodynamic explosions could be at work.

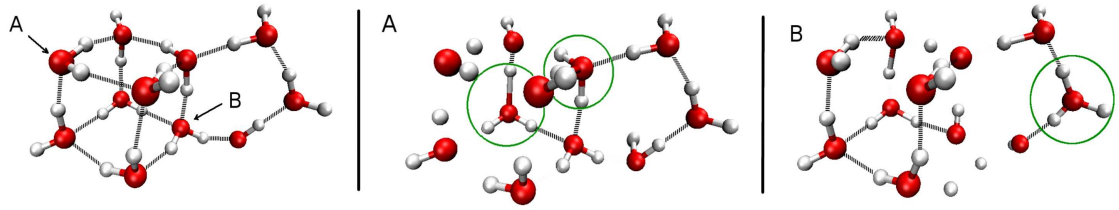


Figure 7.4: Geometry of the $(H_2O)_{11}^{2+}$ cluster obtained from TD-DFT MD simulations at t = 0 (left panel) and t = 12 fs (middle and right panels) when two electrons are removed from the 1b1 orbital of molecules A and B. The circles enclose H_3O^+ units. Oxygen atoms are in red, hydrogen atoms are in grey.

Therefore, future experiments should investigate what kind of size effect and/or excitation mechanisms are responsible for the observed differences. Finally it is also worth mentioning that, as shown by Fig.(7.4), the production of H_3O^+ fragments implies the formation of neutral H, O and OH radicals. This is similar to what happens in liquid water [31] and support the hypothesis of atomic O formation to explain the formation of HO_2 and O_2^- radicals in water radiolysis by swift ions.

Chapter 8

Fragmentation of pyrimidine^{2+} and Auger effect

Abstract *TD-DFT molecular dynamics are extended to investigate fragmentation of doubly-charged gas-phase pyrimidine produced by photoelectron-ion coincidence experiments. Unlikely collision techniques, in which the ionization can occur from very different orbitals, photo-ionization experiments present a major control of the process since there is a much more selectivity of the deposited energy in the target molecule. Thus, combination of both, experimental and theoretical time-dependent tools, is aimed to unambiguously identify the correspondence of the ionized orbital and the mass spectrum signal obtained.*

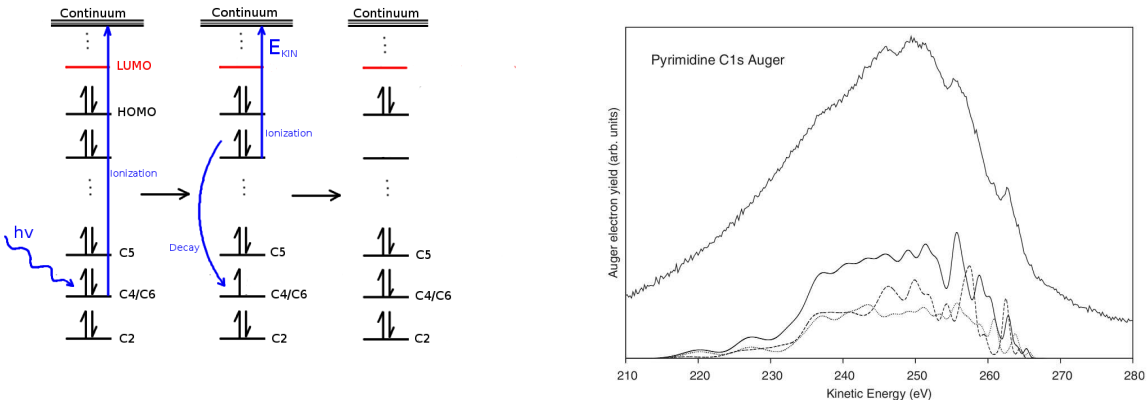
8.1 Introduction

Pyrimidines are an important class of organic molecules mainly due to the fact that the pyrimidine ring forms the base structure of three nucleic acids, i.e. uracil, cytosine and thymine. Furthermore, halogenated pyrimidine bases present an efficient Auger electron emission, and have found applications as radiosensitisers in radiotherapy. For example, it was discovered more than 40 years ago that 5-bromo-deoxyuridine (an analogue of thymidine) radiosensitises cells [247, 248].

During the recent years electron-impact and photoelectron-ion coincidence experiments have provided a great understanding of the electronic structure of such compounds [40–42]. However, so far, experiments are unable to provide insights into the processes that occur at the femtoscale after the electron removal and, therefore, to establish the mechanisms through which the fragments are formed. In the present work we propuse the combination of such techniques with theoretical TD-DFT MD simulations aiming not only to a deep understanding of the fragmentation processes occurred, but also to unambiguously identify the molecular orbital ionized and the fragments produced. This may have important practical consequences since by targeting specific orbitals, one can significantly affect the biological damage.

It is convenient before starting to take a deeper look to the Auger effect and to explain how, by theoretical calculations, the kinetic energies of the emitted electrons can be obtained. In essence, the Auger effect is a physical phenomenon that can be view as a two step process schematically summarized by panel a) of Fig.(8.1): in a first step a core electron is ionized due to X-ray radiation and, second, the k-hole generated is filled by decay of an upper electron. As a consequence of the energy released, a second electron is also ionized (Auger emission) with

certain kinetic energy, leaving the species double charged.



(a) Sketch of the Auger effect

(b) Experimental Auger spectrum of pyrimidine

Figure 8.1: a) Scheme of the Auger effect. b) Solid line, experimental Auger spectrum of pyrimidine taken from Ref. [40].

The kinetic energy of the electrons can be predicted by all-electron calculations using Gaussian09 package [193]. Briefly, we have run an geometry optimization of the neutral pyrimidinine at B3LYP/6-311++G(d,p) level and from the geometry obtained, we have performed two single point calculations over the single and double charged species. These three calculations have allowed as to calculate the first and second vertical ionization energies of pyrimidine. Regarding the Kohn-Sham orbital energies of the neutral molecule, given by Fig.(8.3), we have supposed three core ionization of the k-orbitals centered in C2, C4/C6 and C5 atoms (corresponding to orbitals 3, 4 and 6 from the figure) which are the most probable to occur in electron impact experiment of 1000 eV [40]. The electron emitted is supposed to come from the same orbital of the electron that fills the core vacancy. Its kinetic energy is finally given by the difference between the pyrimidine^{+1} orbital considered and the originally ionized core electron, plus the difference between the the second and first ionization energies. Gathering all the possibilities for each of the k-holes (C2, C4/C6 and C5) we obtain three theoretical spectra shown in Fig.(8.2).

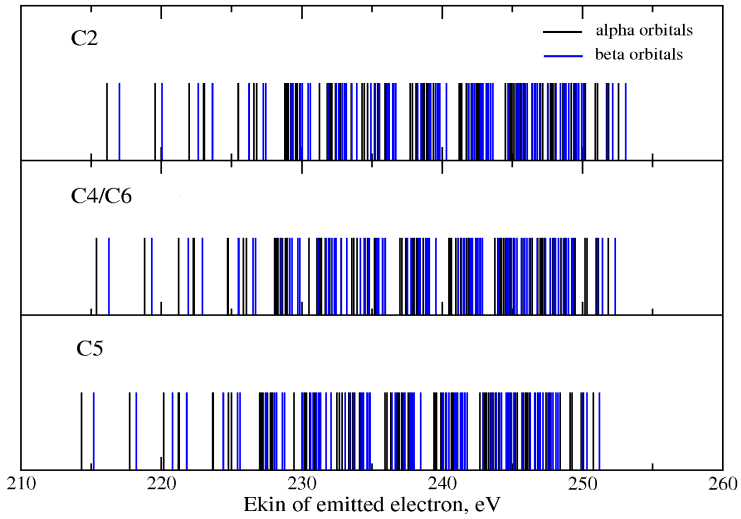


Figure 8.2: Theoretical kinetic energies of the electrons emitted by Auger effect at B3LYP/6-311++G(d,p) level. Black and blue lines represent different spin states.

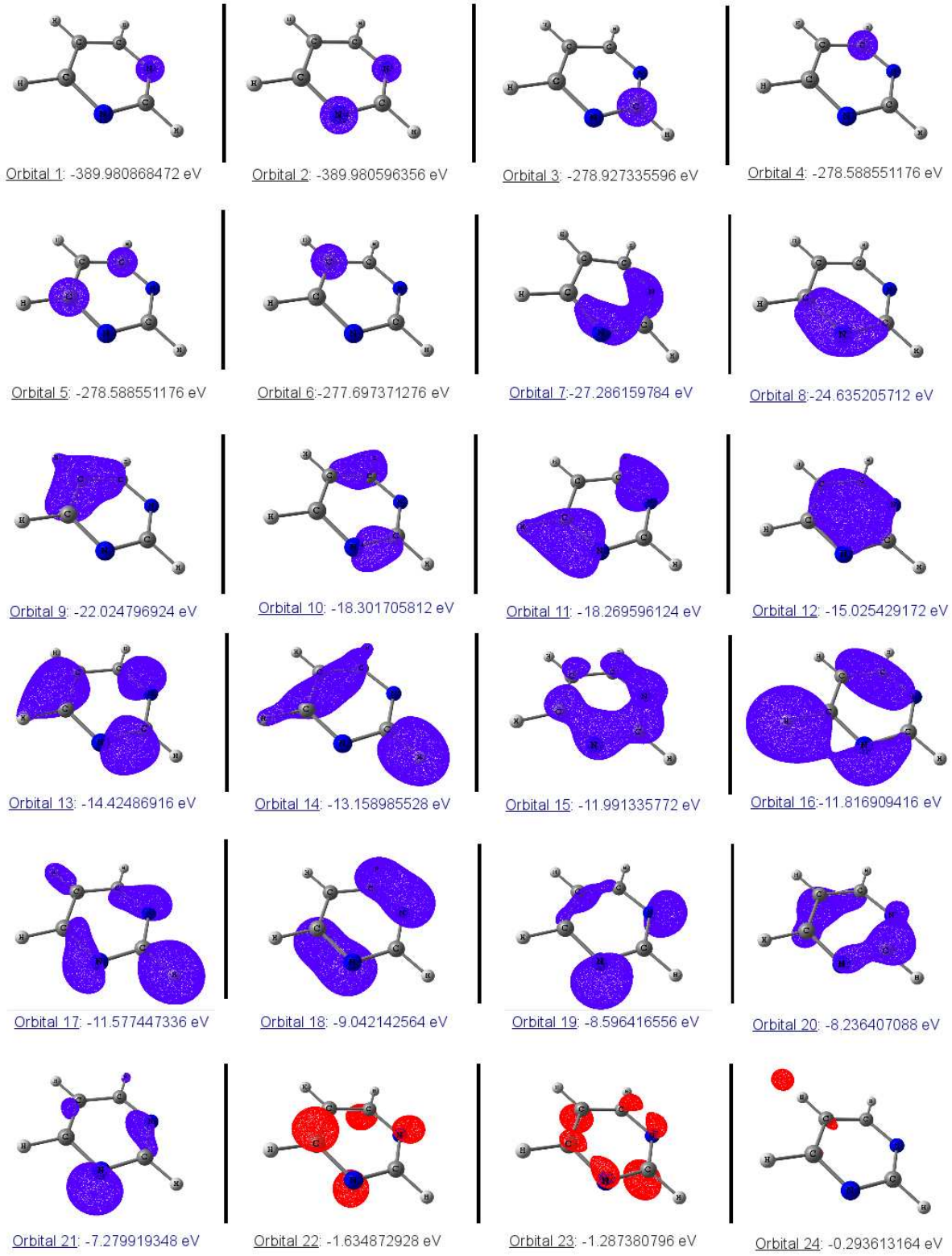


Figure 8.3: Pyrimidine KS orbitals calculated using Gaussian09 at B3LYP/6-311++G(d,p) level. In blue occupied orbitals, in red virtual orbitals. Orbitals covered by pseudopotentials are given in black font letters.

8.2 TD-DFT Molecular Dynamics Simulation

It is worth to notice that predicted kinetic values of Fig.(8.2) are in accordance with the order of magnitude shown by the electron-ion experiment experiments of Fig.(8.2) panel b), in which the signal is given by the composition of the three carbon Auger spectra (C2, C4/C6 and C5 *s* orbitals). Another point of attention is that theoretical values between 210 - 220 eV are spaced out such that make us think in a forthcoming combination with ion-coincidence techniques (signal are barely overlapped). Theory and experiment thus will be able to explicitly addres the origin of the Auger electron produced and the fragments obtained by the Coulombian burst of the dication formed.

The first two lines appeared in the spectra of Fig.(8.2) correspond to the kinetic energies of the Auger electron emitted from the deepest valence orbital (orbital number 7) that represents an electron desity simetrically delocalized between the C2 atom and the two nitrogens. This orbital also corresponds to the deepest KS orbital given by the pseodopotential description of the biomolecule using CPMD package [152], see Fig.(8.4), and its removal can be studied using the TD-DFT MD simulations already used for the study of uracil and small water clusters in gas phase [191, 192].

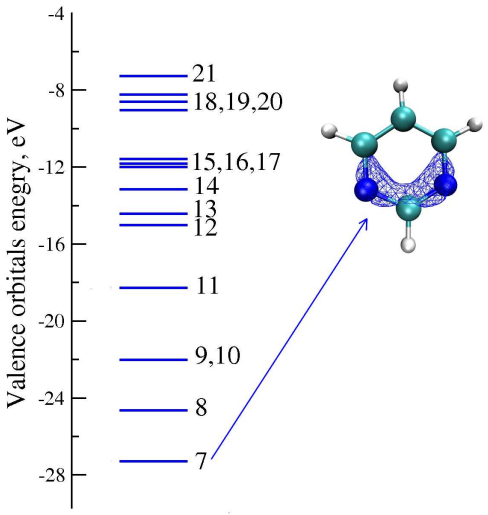


Figure 8.4: Schematic representation of the Kohn-Sham orbital energies obtained by a DFT-base CPMD calculation. The drawing corresponds to the ionized orbital

Thus, we have set-up the initial electron density of the double ion by the removal of the deepest Kohn-Sham orbital of a neutral pyrimidine conformation taken from a thermal equilibration at 300 K. Then the TD-DFT propagation have been extended until $\approx$ 40 fs following the time propagation described in Ref. [33]. We have chosen a step of integration of 0.01 a.u. (0.00024 fs) that assures an energy conservation of $\approx$ 0.1 eV of the total energy of the system.

In the simulation, the wavefunction is represented in a cubic box of 20Å using a cutoff of 70 Ry. Core electrons are replaced by pseudopotentials of the standard Troullier-Martins form [144], giving a total number of 42 electrons described explicitly. The Kleinman-Bylander [207] integration scheme is used for all atoms.

Fig.(8.5) shows the results of TD-DFT MD simulation at a final time of 40 fs. Bonds N1-C6 and C2-N3 are broken almost at the same time $\approx$ 12 fs leading to two fragments of 27 and 53 masses. The bigger one brakes soon after (at 20 fs), leading to 26 and 27 fragments. Thus, at the end of the simulation, three fragments arisen due to the Coulombian explossion: two equivalent

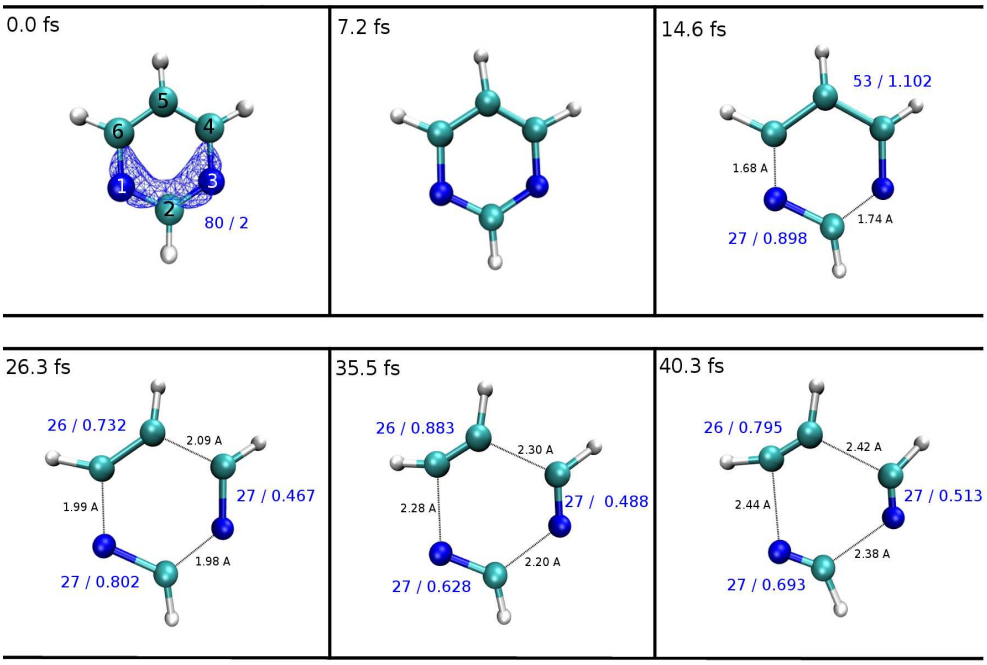


Figure 8.5: Snapshots of the TD-DFT MD trajectory after the KS1 ionization

NCH of 0.513 and 0.693 charged, and C_2H_2 fragment, 0.795 charged. In a mass coincidence spectrum they will lead to a univocal 26/27 signal.

Part IV

Conclusions

Conclusiones

En esta tesis hemos presentado una combinación de Dinámica Molecular TD-DFT/BO para estudiar la etapas iniciales ($\approx$ 100 fs) que suceden tras la doble ionización de biomoléculas tales como uracilo o pirimidina y pequeños agregados de agua. Para ambas metodologías el programa CPMD ha sido usado, empleando la implementación especial de I. Tavernelli y colaboradores para la propagación temporal de las ecuaciones de Kohn-Sham. Los resultados obtenidos han dado lugar a las siguientes conclusiones:

- La simulación de los procesos de fragmentación del ion uracilo^{2+} en fase gas, obtenido mediante la sustracción de dos electrones internos, ha mostrado buen acuerdo con los experimentos de colisión entre moléculas de uracilo y protones de 100 keV. La teoría a sido capaz de predecir señales obtenidas por el espectro de masas en coincidencia, y también ha proporcionado valores correctos para la energía cinética de los fragmentos generados.

- Se ha puesto de manifiesto la relevancia del ambiente químico intramolecular ya que orbitales de misma energía o localizados en enlaces similares, han dado lugar a patrones de fragmentación muy diferentes.

- Los fragmentos observados casi nunca corresponden a los canales de disociación más favorables, hecho relacionado con los efectos dinámicos que tienen lugar en los primeros femtosegundos después de la ionización.

- La doble ionización de uracilo en fase hidratada, generada mediante la sustracción de dos electrones pertenecientes al mismo orbital de Kohn-Sham de la molécula en fase gas, ha producido diferentes patrones de fragmentación. Así, el solvente ha manifestado tomar un papel activo en el proceso de disociación evitando no sólo la movilidad de la carga en el anillo durante los primeros femtosegundos, sino también tomando carga positiva de la biomolécula y llevándola a capas de hidratación más externas.

- La doble ionización de una molécula de agua cercana al uracilo ha dado lugar a la formación de un oxígeno y dos protones. Los protones liberados poseen energías cinética entre 5 y 7.5 eV, dependiendo de si son ionizados de un orbital más o menos profundo.

- Hemos mostrado que la liberación de oxígeno atómico se puede producir tanto por la ionización directa de la molécula de uracilo como la ionización de una molécula de agua del disolvente. Los oxígenos atómicos producidos presentan diferentes cargas y energías cinéticas que influenciarán su posible daño biológico a una escala mayor de tiempo.

- En el estudio de los pequeños agregados de agua, nuestras simulaciones han corroborado la alta movilidad protónica sugerida por el experimento COLTRIMS de colisión. Hemos mostrado que los iones H_3O^+ no son producidos por la molécula de agua que pierde los dos electrones, pero sí de aquellas que capturan los protones liberados de las propias moléculas ionizadas. Por otra parte, fragmentaciones de $[H_2O]_3^{2+}$ han llevado a la formación

de un oxígeno y dos protones, mostrando que el ion H_3O^+ nunca es producido por la fragmenatción del trímero doblemente cargado.

- La combinación de TD-DFT/BO MD y experimentos de coincidencia foto-electrón-ion para el estudio de la fragmentación de iones Pirimidina^{2+} en fase gas proporcionará una asignación inequívoca entre el orbital ionizado y la señal en el espectro de masas producida.

Conclusions

In this thesis we have presented a combination of TD-DFT/BO Molecular Dynamics to study the first stages ($\approx$ 100 fs) that follow the double ionization of biomolecules like uracil or pyrimidine, and small water cluster. We have used the CPMD package with the special implementation of I. Tavernelli et. al for the time-propagation of the Kohn-Sham equations. The results obtained have led to the following conclusions:

- Simulation of the fragmentation process of uracil^{2+} arising from inner shell two-electron vacancies in gas phase, have shown agreement with the experimental collisions of uracil with protons of 100 keV. The theory have been able to predict signals in the mass spectrum coincidence and also to give the proper kinetic energies of the fragments.
- It have been shown the importance of the intramolecular chemical environment since orbitals of similar energy and/or localized in similar bonds have led to very different fragmentation patterns.
- The observed fragments almost never correspond to the energetically most favorable dissociation paths, which is due to dynamical effects occurring in the first few femtoseconds after electron removal.
- Double ionization of uracil in liquid phase, arisen from equivalent Kohn-Sham orbital removal than in gas phase, have produced different fragmentation patterns. The solvent is revealed as an active particpant in the dissociation process since it not only prevents the charge mobility of the biomolecule during the first femtoseconds of dynamics, but also drains positive charge from the uracil to outer hydratation shells.
- Double ionizations of a surrounding water molecule of the uracil steered to the formation of a negative oxygen atom and two protons The energy of the ejected protons are between 5 and 7.5 eV, in accordance with the energy of the ionized orbital.
- We have found that the atomic oxygen release can be produced by the ionization of the uracil or from a water molecule of the solvent. The oxygen atoms produced have different charge and kinetic energy that will influence the biological damage they might cause.
- In the study of small water clusters, our simulations have supported the proton mobility suggested by COLTRIMS experiments. We showed that H_3O^+ ions do not arise from the water molecule that loses the two electrons, but from the water molecules that capture protons ejected from the ionized ones. On the other hand, fragmentations of $[H_2O]_3^{2+}$ led to one oxygen atom and two protons, showing that H_3O^+ ion is never produced by the Coulomb explosion of the double charged trimer.
- Combination of TD-DFT/BO MD and photoelectron-ion coincidence experiments to study the fragmentation of Pyrimidine^{2+} ion are aimed to unambiguously identify the correspondence between the orbital ionized and the mass spectra signals obtained.

Part V

Appendixes

Appendix A

Appendix

A.1 Ehrenfest theorem

The Ehrenfest theorem explains how the mean value of an observable $< A >$ varies with time:

$$\frac{d\langle A\rangle}{dt} = \left\langle \frac{\partial A}{\partial t} \right\rangle + \frac{i}{\hbar}\langle [H, A]\rangle \tag{A.1}$$

where $\langle [H, A]\rangle$ is the average of the Hamiltonian conmutator defined as $< HA - AH >$ The proof starts by considering a time time-dependent wavefunction, $|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle$, that satisfy the TDSE (2.1). Then, the variation of time of the average observable $< A >$ is guiven by:

$$\frac{d\langle A\rangle}{dt} = \langle \frac{\partial}{\partial t}\Phi(\mathbf{r}, \mathbf{R}, t)|\hat{A}|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle + \langle \Phi(\mathbf{r}, \mathbf{R}, t)|\frac{\partial \hat{A}}{\partial t}|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle + \\ + \langle \Phi(\mathbf{r}, \mathbf{R}, t)|\hat{A}|\frac{\partial}{\partial t}\Phi(\mathbf{r}, \mathbf{R}, t)\rangle \tag{A.2}$$

where we can identify the second term on the right as the partial derivate of the mean value of $\hat{A}$ respect to time. Considering now the postulated time variations of the wavefunction:

$$\langle \frac{\partial}{\partial t}\Phi(\mathbf{r}, \mathbf{R}, t)| = -\frac{1}{i\hbar}(\hat{H}|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle)^* \quad and \quad |\frac{\partial}{\partial t}\Phi(\mathbf{r}, \mathbf{R}, t)\rangle = \frac{1}{i\hbar}\hat{H}|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle \tag{A.3}$$

First and third term of right-hand side of Eq. (A.2) can be written as:

$$\langle \frac{\partial}{\partial t}\Phi(\mathbf{r}, \mathbf{R}, t)|\hat{A}|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle + \langle \Phi(\mathbf{r}, \mathbf{R}, t)|\hat{A}|\frac{\partial}{\partial t}\Phi(\mathbf{r}, \mathbf{R}, t)\rangle = \\ = -\frac{1}{i\hbar}\langle \hat{H}\Phi(\mathbf{r}, \mathbf{R}, t)|\hat{A}|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle + \frac{1}{i\hbar}\langle \Phi(\mathbf{r}, \mathbf{R}, t)|\hat{A}\hat{H}|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle \tag{A.4}$$

and considering the hermeticity of both operators, Eq.(A.4) leads to:

$$\langle \frac{\partial}{\partial t}\Phi(\mathbf{r}, \mathbf{R}, t)|\hat{A}|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle + \langle \Phi(\mathbf{r}, \mathbf{R}, t)|\hat{A}|\frac{\partial}{\partial t}\Phi(\mathbf{r}, \mathbf{R}, t)\rangle = \\ = \frac{i}{\hbar}\langle \Phi(\mathbf{r}, \mathbf{R}, t)|\hat{H}\hat{A}|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle - \frac{i}{\hbar}\langle \Phi(\mathbf{r}, \mathbf{R}, t)|\hat{A}\hat{H}|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle \tag{A.5}$$

corresponding with the definition of the conmutator $\langle [H, A]\rangle$.

A.2 The BO antsaz

Substituting the Born-Oppenheimer ansatz of Eq.(2.43) into the equation Time-dependent Schrödinger Equation (2.1), we get:

$$i\hbar\frac{\partial}{\partial t}\sum_{l=0}^{\infty}|\Psi_l(\mathbf{r};\mathbf{R})\rangle|\chi_l(\mathbf{R},t)\rangle = \hat{H}\sum_{l=0}^{\infty}|\Psi_l(\mathbf{r};\mathbf{R})\rangle|\chi_l(\mathbf{R},t)\rangle \tag{A.6}$$

From the left part of the Eq.(A.6):

$$\begin{aligned} &i\hbar\frac{\partial}{\partial t}\sum_{l=0}^{\infty}|\Psi_l(\mathbf{r};\mathbf{R})\rangle|\chi_l(\mathbf{R},t)\rangle = \\ &= i\hbar\sum_{l=0}^{\infty}\frac{\partial|\Psi_l(\mathbf{r};\mathbf{R})\rangle}{\partial t}|\chi_l(\mathbf{R},t)\rangle + i\hbar\sum_{l=0}^{\infty}|\Psi_l(\mathbf{r};\mathbf{R})\rangle\frac{\partial|\chi_l(\mathbf{R},t)\rangle}{\partial t} \end{aligned} \tag{A.7}$$

where the first term is equal to zero. Thus, replacing Eq.(A.7) and Eq.(2.3) into Eq.(A.6):

$$\begin{aligned} &i\hbar\sum_{l=0}^{\infty}|\Psi_l(\mathbf{r};\mathbf{R})\rangle\frac{\partial|\chi_l(\mathbf{R},t)\rangle}{\partial t} = \\ &= \hat{H}_e(\mathbf{r},\mathbf{R})\sum_{l=0}^{\infty}|\Psi_l(\mathbf{r};\mathbf{R})\rangle|\chi_l(\mathbf{R},t)\rangle + \hat{T}_N(\mathbf{R})\sum_{l=0}^{\infty}|\Psi_l(\mathbf{r};\mathbf{R})\rangle|\chi_l(\mathbf{R},t)\rangle \end{aligned} \tag{A.8}$$

Multiplying both sides of Eq.(A.8) by $\langle\Psi_k(\mathbf{r};\mathbf{R})|$:

$$\begin{aligned} &i\hbar\sum_{l=0}^{\infty}\langle\Psi_k(\mathbf{r};\mathbf{R})|\Psi_l(\mathbf{r};\mathbf{R})\rangle\frac{\partial|\chi_l(\mathbf{R},t)\rangle}{\partial t} = \sum_{l=0}^{\infty}\langle\Psi_k(\mathbf{r};\mathbf{R})|E_l(\mathbf{R})|\Psi_l(\mathbf{r};\mathbf{R})\rangle|\chi_l(\mathbf{R},t)\rangle + \\ &+\sum_{l=0}^{\infty}\langle\Psi_k(\mathbf{r};\mathbf{R})|\hat{T}_N(\mathbf{R})|\Psi_l(\mathbf{r};\mathbf{R})\rangle|\chi_l(\mathbf{R},t)\rangle \end{aligned} \tag{A.9}$$

The second term in the right side of Eq.(A.9) is equal to:

$$\begin{aligned} &\sum_{l=0}^{\infty}\langle\Psi_k(\mathbf{r};\mathbf{R})|\hat{T}_N(\mathbf{R})|\Psi_l(\mathbf{r};\mathbf{R})\rangle|\chi_l(\mathbf{R},t)\rangle = \\ &= \sum_{l=0}^{\infty}\langle\Psi_k(\mathbf{r};\mathbf{R})|\chi_l(\mathbf{R},t)\rangle\hat{T}_N(\mathbf{R})|\Psi_l(\mathbf{r};\mathbf{R})\rangle + \sum_{l=0}^{\infty}\langle\Psi_k(\mathbf{r};\mathbf{R})|\Psi_l(\mathbf{r};\mathbf{R})\rangle\hat{T}_N(\mathbf{R})|\chi_l(\mathbf{R},t)\rangle - \\ &-\sum_{l=0}^{\infty}\sum_{\alpha=1}^{N}\frac{\hbar^2}{M_\alpha}\langle\Psi_k(\mathbf{r};\mathbf{R})|\nabla_\alpha|\Psi_l(\mathbf{r};\mathbf{R})\rangle\nabla_\alpha|\chi_l(\mathbf{R},t)\rangle \end{aligned} \tag{A.10}$$

Knowing this expression, and canceling the summations, Eq.(A.8) can be expresed as:

$$\begin{aligned} i\hbar\frac{\partial|\chi_k(\mathbf{R},t)\rangle}{\partial t} &= \\ &= [\hat{T}_N(\mathbf{R}) + E_k(\mathbf{R})]|\chi_k(\mathbf{R},t)\rangle + \\ +\sum_{l=0}^{\infty}&\left(\langle\Psi_k(\mathbf{r};\mathbf{R})|\hat{T}_N(\mathbf{R}|\Psi_l(\mathbf{r};\mathbf{R})\rangle - \sum_{\alpha=1}^{N}\frac{\hbar^2}{M_\alpha}\langle\Psi_k(\mathbf{r};\mathbf{R})|\nabla_\alpha|\Psi_l(\mathbf{r};\mathbf{R})\rangle\nabla_\alpha\right)|\chi_l(\mathbf{R},t)\rangle \end{aligned} \tag{A.11}$$

where the terms in parenthesis correspond to the definition of coupling operator defined in Eq.(2.45).

A.3 Semiclassical equations

Substitutinq Eq.(2.52) into the TISE Eq.(2.51) we obtain:

$$[\hat{T}_N(\mathbf{R}) + E_k(\mathbf{R})]A_k(\mathbf{R},t)\exp(\frac{i}{\hbar}S_k(\mathbf{R},t)) = i\hbar\frac{\partial}{\partial t}A_k(\mathbf{R},t)\exp(\frac{i}{\hbar}S_k(\mathbf{R},t)) \tag{A.12}$$

That can be written as:

$$\begin{aligned}
&-\sum_{\alpha=1}^{N}\frac{\hbar^2}{2M_\alpha}\nabla_\alpha^2 A_k(\mathbf{R},t)\exp(\frac{i}{\hbar}S_k(\mathbf{R},t)) + E_k(\mathbf{R})A_k(\mathbf{R},t)\exp(\frac{i}{\hbar}S_k(\mathbf{R},t)) = \\
&= i\hbar\exp(\frac{i}{\hbar}S_k(\mathbf{R},t))\frac{\partial A_k(\mathbf{R},t)}{\partial t} - A_k(\mathbf{R},t)\exp(\frac{i}{\hbar}S_k(\mathbf{R},t))\frac{\partial S_k(\mathbf{R},t)}{\partial t}
\end{aligned} \tag{A.13}$$

Expanding the left term of the equation above:

$$\begin{aligned}
&-\sum_{\alpha=1}^{N}\frac{\hbar^2}{2M_\alpha}A_k(\mathbf{R},t)\exp(\frac{i}{\hbar}S_k(\mathbf{R},t)) + E_k(\mathbf{R})A_k(\mathbf{R},t)\exp(\frac{i}{\hbar}S_k(\mathbf{R},t)) = \\
&= -\sum_{\alpha=1}^{N}\frac{\hbar^2}{2M_\alpha}\exp(\frac{i}{\hbar}S_k(\mathbf{R},t))\nabla_\alpha^2 A_k(\mathbf{R},t) - \sum_{\alpha=1}^{N}\frac{\hbar^2}{2M_\alpha}A_k(\mathbf{R},t)\nabla_\alpha^2\exp(\frac{i}{\hbar}S_k(\mathbf{R},t)) - \\
&- \sum_{\alpha=1}^{N}\frac{\hbar^2}{M_\alpha}\nabla_\alpha A_k(\mathbf{R},t)\nabla_\alpha\exp(\frac{i}{\hbar}S_k(\mathbf{R},t)) + E_k(\mathbf{R})]A_k(\mathbf{R},t)\exp(\frac{i}{\hbar}S_k(\mathbf{R},t))
\end{aligned} \tag{A.14}$$

Regarding that:

$$\nabla_\alpha\exp(\frac{i}{\hbar}S_k(\mathbf{R},t)) = \frac{i}{\hbar}\nabla_\alpha S_k(\mathbf{R},t))\exp(\frac{i}{\hbar}S_k(\mathbf{R},t)) \tag{A.15}$$

and also the second derivative:

$$\begin{aligned}
&\nabla_\alpha^2\exp(\frac{i}{\hbar}S_k(\mathbf{R},t)) = \nabla_\alpha(\frac{i}{\hbar}\nabla_\alpha S_k(\mathbf{R},t))\exp(\frac{i}{\hbar}S_k(\mathbf{R},t))) = \\
&= \frac{i}{\hbar}\exp(\frac{i}{\hbar}S_k(\mathbf{R},t))\nabla_\alpha^2 S - \frac{1}{\hbar^2}\exp(\frac{i}{\hbar}S_k(\mathbf{R},t))(\nabla_\alpha S_k(\mathbf{R},t)))^2
\end{aligned} \tag{A.16}$$

and dividing the expresion Eq.(A.14) by Eq.(2.52), we obtain:

$$\begin{aligned}
&\sum_{\alpha=1}^{N}\frac{\hbar^2}{2M_\alpha}A_k(\mathbf{R},t)\exp(\frac{i}{\hbar}S_k(\mathbf{R},t)) + E_k(\mathbf{R})A_k(\mathbf{R},t)\exp(\frac{i}{\hbar}S_k(\mathbf{R},t)) = \\
&= -\sum_{\alpha=1}^{N}\frac{\hbar^2}{2M_\alpha}\frac{\nabla_\alpha^2 A_k(\mathbf{R},t)}{A_k(\mathbf{R},t)} - \sum_{\alpha=1}^{N}\frac{i\hbar}{2M_\alpha}\nabla_\alpha^2 S_k(\mathbf{R},t) + \sum_{\alpha=1}^{N}\frac{1}{2M_\alpha}(\nabla_\alpha S_k(\mathbf{R},t))^2 - \\
&- \sum_{\alpha=1}^{N}\frac{i\hbar}{M_\alpha}\frac{\nabla_\alpha A_k(\mathbf{R},t)\nabla_\alpha S_k(\mathbf{R},t)}{A_k(\mathbf{R},t)} + E_k(\mathbf{R})
\end{aligned} \tag{A.17}$$

Finally, matching this expression with the right hand side of Eq.(A.13), previously divided by Eq.(2.52), we get:

$$\begin{aligned}
&-\sum_{\alpha=1}^{N}\frac{\hbar^2}{2M_\alpha}\frac{\nabla_\alpha^2 A_k(\mathbf{R},t)}{A_k(\mathbf{R},t)} - \sum_{\alpha=1}^{N}\frac{i\hbar}{2M_\alpha}\nabla_\alpha^2 S_k(\mathbf{R},t) + \sum_{\alpha=1}^{N}\frac{1}{2M_\alpha}(\nabla_\alpha S_k(\mathbf{R},t))^2 - \\
&- \sum_{\alpha=1}^{N}\frac{i\hbar}{M_\alpha}\frac{\nabla_\alpha A_k(\mathbf{R},t)\nabla_\alpha S_k(\mathbf{R},t)}{A_k(\mathbf{R},t)} + E_k(\mathbf{R}) = i\hbar\frac{1}{A_k(\mathbf{R},t)}\frac{\partial A_k(\mathbf{R},t)}{\partial t} - \frac{\partial S_k(\mathbf{R},t)}{\partial t}
\end{aligned}$$

where real and imaginary terms can be separated, leading straightforwardly to Eq.(2.53) and Eq.(2.54).

A.4 Fundamentals of Hartree-Fock Aproximation

A.4.1 Variational theorem

Assuming that we know the exact solutions to the Schrödinger equation, being E_0 the lowest energy value:

$$\hat{H}|\Psi(\mathbf{r},\mathbf{R})_k\rangle = E_k|\Psi(\mathbf{r},\mathbf{R})_k\rangle \qquad k = 0,1,2,...,\infty \tag{A.18}$$

If $|\phi(\mathbf{r},\mathbf{R})\rangle$ is any well-behaved function of the same variables as $|\Psi_k(\mathbf{r},\mathbf{R})\rangle$ and satisfies the same boundary conditions, then:

$$\xi = \frac{\langle\phi(\mathbf{r},\mathbf{R})|\hat{H}|\phi(\mathbf{r},\mathbf{R})\rangle}{\langle\phi(\mathbf{r},\mathbf{R})|\phi(\mathbf{r},\mathbf{R})\rangle} \geq E_0 \tag{A.19}$$

Thus, the variational theorem establish an upper energy limit for the ground state of the system, which allow us to successively compare our energy guess in order to obtain the lowest possible value of ξ.

A.4.2 Slater determinant

The wave function that represents a system formed by n equal non-interacting electrons is:

$$|\Psi(\chi_1,\chi_2,\chi_3,...,\chi_n)\rangle = \prod_{i=1}^{n}|\chi_i(\mathbf{x})\rangle \tag{A.20}$$

where functions $|\chi_i(\mathbf{x})\rangle$ are called spin-orbitals and each one represents a non-interacting electron. In turn, they are made up of the product of two functions, one to describe the spatial distribution $|\psi(\mathbf{r})\rangle$, and other, $|\alpha(\omega)\rangle$ or $|\beta(\omega)\rangle$, to describe its spin:

$$|\chi_i(\mathbf{x})\rangle = \begin{cases} |\psi(\mathbf{r})\rangle|\alpha(\omega)\rangle \\ |\psi(\mathbf{r})\rangle|\beta(\omega)\rangle \end{cases} \tag{A.21}$$

The *antisymmetry principle of Pauli* [249] establishes that a many-electron wave function must be antisymmetric with respect to the interchange of the coordinates (both space and spin) of any two electrons:

$$|\Psi(\chi_1,\chi_2,\chi_3,...,\chi_n)\rangle = -|\Psi(\chi_1,\chi_3,\chi_2,...,\chi_n)\rangle \tag{A.22}$$

Thus, the antisymmetry condition must be imposed on many-electron wave function (A.20) to be solution of the Schrödinger equation (2.40):

$$|\Psi_e\rangle = (N!)^{-1/2}\sum_{r=1}^{N!}\delta_r\hat{P}_r|\Psi(\chi_1,\chi_2,\chi_3,...,\chi_n)\rangle \tag{A.23}$$

where the operator $\hat{P}_r$ is any one of the N! operators, including the identity operator, that permute a given order of particles to another order. The summation is taken over all N! permutation operators. The quantity δ_r is +1 or -1 if the permutation operator $\hat{P}_r$ involves the exchange of an even or odd number of particle pairs. The factor $(N!)^{-1/2}$ normalizes $|\Psi_e\rangle$ if $|\Psi(\chi_1,\chi_2,\chi_3,...,\chi_n)\rangle$ is normalized.
The equation (A.23) corresponds to the Slater Determinant, usually denoted as $|\chi_1\chi_2\chi_3...\chi_n|$

A.4.3 Hartree-Fock method

The Hartree-Fock method or Self-Consistent Field (SCF) method [250] represents an approximate variational method to determine the ground-state wave function and ground-state energy of a quantum many-electron system where the wave function is assumed to be a single Slater determinant $|\Psi_e\rangle$ of n spin-orbitals (A.23).

The Hartree-Fock approximation constitutes the simplest *ab-initio* approximation to solve the time-independent Schrödinger equation and is the starting point of more accurate methods that include higher electronic correlation. Breafly, it consists in minimize the variational integral with the constraint that the spin-orbitals remain orthonormal:

$$L[\{\chi_i\}] = \xi[\{\chi_i\}] + \sum_{i=1}^{n}\sum_{j=1}^{n} l_{ij}(\langle\chi_i(\mathbf{1})|\chi_j(\mathbf{1})\rangle - \delta ij) \tag{A.24}$$

Process in which the iterative resulution of the "eigenvalue equations" known as *Hartree-Fock equations*:

$$\begin{array}{l} \hat{f}(\mathbf{1})|\chi_1'(\mathbf{1})\rangle = \epsilon_1|\chi_1'(\mathbf{1})\rangle \\ \hat{f}(\mathbf{1})|\chi_2'(\mathbf{1})\rangle = \epsilon_2|\chi_2'(\mathbf{1})\rangle \\ \quad\vdots \quad\quad \vdots \quad\quad \vdots \\ \hat{f}(\mathbf{1})|\chi_n'(\mathbf{1})\rangle = \epsilon_n|\chi_n'(\mathbf{1})\rangle \end{array} \tag{A.25}$$

where ϵ_i term represents each time the energy of an electron in the orbital-orbital $|\chi_i\rangle$ subject to interaction with all the other electrons:

$$\epsilon_i = H_{ii}(\mathbf{1}) + \sum_{j=1}^{n}(J_{ij}(\mathbf{1}) - K_{ij}(\mathbf{1})) \tag{A.26}$$

Thus, the complicated multielectron problem is replaced by many monoelectronic problems in which the electrons are described as independent particles interacting with each other through a mean-averaged repulsion potential.

Appendix B

Appendix

B.1 Fragmentation channels

Table B.1: Estimated energies for uracil^{2+} ions generated by removing two electrons from the three deepest orbitals and HOMO orbital. For each case, the energy is approximately given by subtracting two times the corresponding Gaussian-BLYP KS-orbital energy to the CCSD(T) energy of neutral uracil.

Uracil^{2+}	E (a.u.)	E (eV)
Hole in KS1	-412.20308	-11216.61601
Hole in KS2	-412.26756	-11218.37062
Hole in KS3	-412.38146	-11221.47002
Hole in Valence	-413.72207	-11257.94989

Table B.2: Different fragmentation channels with total energies in eV. For each orbital, relative energies (E_{rel}) are referred to uracil^{2+} values of Table(B.1). In blue, the channels obtained by the simulations, in red the most stable path.

	Fragments	Peak	E (eV)	E_{rel} KS1 (eV)	E_{rel} KS2 (eV)	E_{rel} KS3 (eV)	E_{rel} HOMO (eV)
1	$HNCCH_2^{1+} + OCNH^{1+} + CO$	41/43	-11246.77234	-30.16	-28.40	-25.30	11.18
2	$HNCCH_2^{1+} + OCNH + CO^{1+}$	41/28	-11244.36876	-27.75	-26.00	-22.90	13.58
3	$HNCCH_2 + OCNH^{1+} + CO^{1+}$	43/28	-11241.53764	-24.92	-23.17	-20.07	16.41
4	$HCCH^{1+} + CNH^{1+} + O(NH)C + O$	26/27	-11232.92048	-16.30	-14.55	-11.45	25.03
5	$HCCH^{1+} + CNH + O(NH)C^{1+} + O$	26/43	-11233.80076	-17.18	-15.43	-12.33	24.15
6	$HCCH^{1+} + CNH + O(NH)C + O^{1+}$	26/16	-11231.46252	-14.85	-13.09	-9.99	26.49
7	$HCCH + CNH^{1+} + O(NH)C^{1+} + O$	27/43	-11233.16764	-16.55	-14.80	-11.70	24.78
8	$HCCH + CNH^{1+} + O(NH)C + O^{1+}$	27/16	-11230.82940	-14.21	-12.46	-9.36	27.12
9	$HCCH + CNH + O(NH)C^{1+} + O^{1+}$	43/16	-11231.70968	-15.09	-13.34	-10.24	26.24
10	$C_2H_2^{1+} + OCNH^{1+} + OCNH$	26/43	-11243.72892	-27.11	-25.36	-22.26	14.22
11	$C_2H_2 + OCNH^{1+} + OCNH^{1+}$	43/43	-11243.50313	-26.89	-25.13	-22.03	14.45
12	$HNCH_2CO^{1+} + OCNH^{1+}$	69/43	-11258.20567	-41.59	-39.84	-36.74	-0.26
13	$HNCH_2CO^{2+} + OCNH$	–	-11244.60809	-27.99	-26.24	-23.14	13.34
14	$HCCH^{1+} + CO^{1+} + OCNH + NH$	26/28	-11237.55032	-20.93	-19.18	-16.08	20.40
15	$HCCH^{1+} + CO + OCNH^{1+} + NH$	26/43	-11239.95389	-23.34	-21.58	-18.48	18.00
16	$HCCH^{1+} + CO + OCNH + NH^{1+}$	26/15	-11238.06003	-21.44	-19.69	-16.59	19.89
17	$HCCH + CO^{1+} + OCNH^{1+} + NH$	28/43	-11237.32454	-20.71	-18.95	-15.85	20.63
18	$HCCH + CO^{1+} + OCNH + NH^{1+}$	28/15	-11235.43068	-18.81	-17.06	-13.96	22.52
19	$HCCH + CO + OCNH^{1+} + NH^{1+}$	43/15	-11237.83425	-21.22	-19.46	-16.36	20.12
20	$HNC_3H_2^{1+} + OCNH^{1+} + O$	53/43	-11241.50292	-24.89	-23.13	-20.03	16.45
21	$HNC_3H_2^{1+} + OCNH + O^{1+}$	53/16	-11239.63762	-23.02	-21.27	-18.17	18.31
22	$HNC_3H_2 + OCNH^{1+} + O^{1+}$	43/16	-11236.69055	-20.07	-18.32	-15.22	21.26
23	$HNCCH_2CO^{1+} + CNH^{1+} + O$	69/27	-11240.50512	-23.89	-22.13	-19.04	17.44
24	$HNCCH_2CO^{1+} + CNH + O^{1+}$	69/16	-11239.04716	-22.43	-20.68	-17.58	18.90

B.2 Dissociation barriers

We have stimated the order of magnitude of the dissociation barriers of uracil^{2+} by propossing two fragmentations similar to the ones obtained in KS1 and KS3 dynamics. The calculations have been done using Gaussian09 at BLYP/6-311+G(d,p) level and transition states have been obtained by QST3 procedure followed by a frequency calculation. For neutral fragment loss, the energy of the process is stimated by the binding energy.

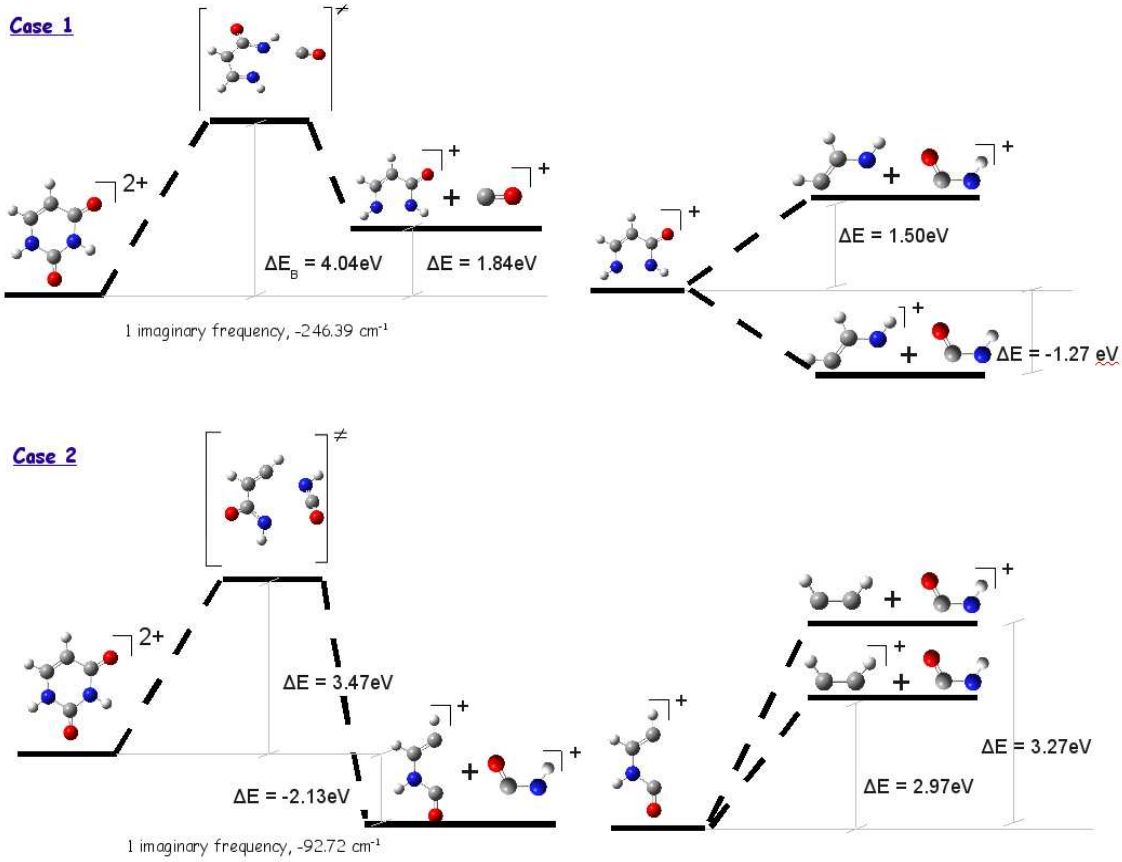
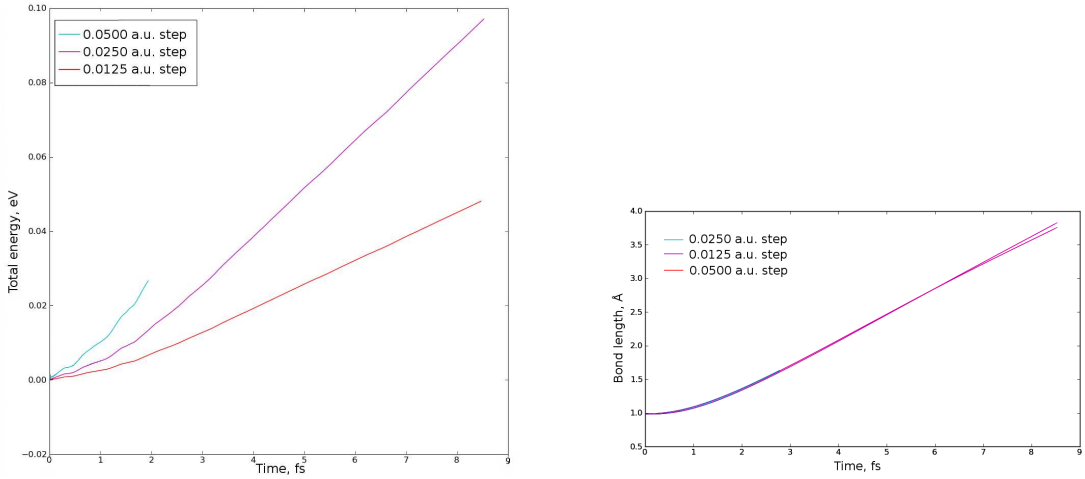


Figure B.1: Calculations of two transitions states and different binding energies for two different fragmentations of uracil^{2+}

B.3 Time-step tests in liquid water

Several time-steps for the time-dependent propagation have been tested by ionizing a water molecule in liquid. As it is summarized in the figures below, although the energy is worse conserved with a 0.05 a.u. time-step (left panel), for smaller steps, at short times of dynamics, no big changes have been detected in the O-H distances of the ionized water molecule (right panel).



(a) Total energy

(b) O-H distance

Figure B.2: Total energy conservation (a) and O-H distance evolution (b) using different time steps for a TD-DFT dynamics simulation of a double ionization in liquid water

Appendix C

List of Publications

- Journal: Phys. Rev. Lett. submitted

  Authors: P. López-Tarifa M.A. Hervé du Penhoat, R. Vuilleumier, I. Tavernelli, M.-P. Gaigeot, J.P. Champeaux, A. Le Padellec, M. Alcamí, P. Moretto-Capelle, F. Martín, and M.-F. Politis

  Title: "Ultrafast non-adiabatic fragmentation dynamics of doubly charged uracil in gas phase"

- Journal: Mutation Research-Reviews in Mutation Research (2010)

  Authors: M-P. Gaigeot, P. López-Tarifa, F. Martín, M. Alcamí, R. Vuilleumier, I. Tavernelli, M. A. Hervé du Penhoat and M-F. Politis

  Volume/Pages: 704, 45

  Title: "Theoretical Investigation of the Ultrafast Dissociation of Ionised Biomolecules Immersed in Water: Direct and Indirect Effects"

- Journal: J. Phys. B: At. Mol. Opt. Phys. (2009)

  Authors: L. Adoui, A. Cassimi, B. Gervais, L. Guillaume, S. Legendre, M. Tarisien, P. López-Tarifa, M-F. Politis, M. A. Hervé du Penhoat, R. Vuilleumier, M-P. Gaigeot, I. Tavernelli, M. Alcamí , and F. Martín

  Volume/Pages: 42, 075101

  Title: "Ionization and fragmentation of water clusters by fast highly charged ions"

Agradecimientos

Esta tesis se ha sido relidazada entre el Departamento de Química de la Universidad Autónoma de Madrid (UAM) y el Laboratoire Analyse et Modélisation pour la Biologie et l'Environnementel (LAMBE) de la Université Evry Val d'Essonne durante los años 2007-2010 bajo la co-dirección de los Profesores Prof. Manuel Alcamí Pertejo y Marie-Pierre Gaigeot, en colaboración con los Profesores Marie-Franoise Politis (LAMBE), Fernando Martín (UAM), Marie-Anne Hervè du Penhoat (Institut de Minéralogie et de Physique des Milieux Condensés, Paris -France), Rodolphe Villeumier (École Normale Supérieure, Paris -France) e Ivano Tavernelli (École Polytechnique Fédérale de Lausanne -Switzerland).

A todos vosotros, en conjunto, me gustaría agradecer todo vuestro esfuerzo, vuestra dedicación y ayuda, tanto en las estancias en vuestros laboratorios, como fuera de ellos. Gracias por vuestros mails, consejos y apoyo infinito. Gracias en definitiva por haberme dado la increíble oportunidad de poder trabajar a vuestro lado.

También agradecer a todos los becarios, postdocs y profesores del departamento C-IX, gracias por vuestra incansable vuestro apoyo y sabios consejos.

Gracias también a mis amigos y familiares por haber soportado la otra cara de un doctorado. Gracias Ele por tu paciencia, apoyo infinito y serenidad. Gracias a mis padres por vuestra completa dedicación, por haberme transmitido vuestro esfuerzo, humildad e ilusión.

List of Figures

List of Tables

References

[1] J. Ward, Prog. Nucleic Acid Res. Mol. Biol. **35**, 95 (1988).

[2] G. A. Jeffrey and W. Saenger, *Hydrogen Bonding in Biological Structures* (Springer-Verlag, New York, 1991).

[3] S. Gregoli, S. M. Olast, and A. Bertinchamps, Radiat. Res. **89**, 238 (1982).

[4] D. Becker and M. Sevilla, Adv. Radiat. Biol. **17**, 121 (1993).

[5] T. L. Vere, D. Becker, and M. D. Sevilla, Radiat. Res. **145**, 673 (1996).

[6] M. Debije, M. Strickler, and A. Bernhard, Radiat. Res. **154**, 163 (2000).

[7] Y. Gauduel, *Ultrafast Dynamics of Chemical Systems* (Kluwer Publisher, ADDRESS, 1994).

[8] B. Brozek-Pluska *et al.*, Radiat. Phys. Chem **72**, 149 (2005).

[9] Y. Gauduel *et al.*, Biochemistry **27**, 2509 (1988).

[10] Y. Gauduel, H. Gelabert, and F. Guilloud, J. Am. Chem. Soc. **112**, 5082 (2000).

[11] W. Chun-Rong, J. Nguyen, and L. Qing-Bin, J. Am. Chem. Soc. **113**, 11320 (2009).

[12] F. Ban, K. Rankin, J. Gauld, and R. Boyd, Theor. Chem. Acc. **108**, 11 (2002).

[13] S. Naumov and C. von Sonntag, Radiat. Res. **169**, 355 (2008).

[14] N. Jena and P. C. Mishra, J. Phys. Chem. B. **109**, 14205 (2005).

[15] A. Kumar and M. Sevilla, *Radiation effects on DNA: theoretical investigations of electron hole and excitation pathways to DNA damage* (Springer-Verlag, Berlin, 2008).

[16] X. Li and M. Sevilla, Adv. Quant. Chem. **52**, 59 (2007).

[17] R. H. D. Lyngdoh and H. S. III, Acc. Chem. Res. **42**, 563 (2009).

[18] C. J. Mundy, M. E. Colvin, and A. A. Quong, J. Phys. Chem. A **106**, 10063 (2002).

[19] Y. Wu, C. Mundy, M. Colvin, and R. Car, J. Phys. Chem. A **108**, 2922 (2004).

[20] C. Champion, *Ph.D. thesis* (PUBLISHER, University Paris7, France, 1999).

[21] N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions*, 2nd ed. (Oxford University Press, Oxford at the Clarendon Press, 1952).

[22] M. R. C. M. Dowell and J. Coleman, *The Theory of Atomic Collisions* (Introduction to the Theory of the Ion-Atom Collision, North Holland, Amsterdam, 1970).

[23] F. Alvarado *et al.*, Chem. Phys. Chem. **9**, 1254 (2008).

[24] P. Moretto-Capelle and A. L. Padellec, Phys. Rev. A **74**, 062705 (2006).

[25] T. Schlatholter, R. Hoekstra, and R. Morgenstern, Int. J. Mass Spectrom. **233**, 173 (2004).

[26] A. Chetioui *et al.*, Int. J. Radiat. Biol. **65**, 511 (1994).

[27] B. Fayard *et al.*, Radiat. Res. 157 (2002) 128140. **157**, 128 (2002).

[28] A. Eschenbrenner *et al.*, Int. J. Radiat. Biol. **83**, 687 (2007).

[29] B. Gervais, M. Beuve, G. Olivera, and M. Galassi, Radiat. Phys. Chem. **75**, 493 (2006).

[30] G. Baldacchino *et al.*, Nucl. Instr. Meth. Phys. Res. B. **146**, 528 (1998).

[31] M.-P. Gaigeot *et al.*, Phys. B: At. Mol. Opt. Phys. **40**, 1 (2007).

[32] G. Baldacchino *et al.*, Radiat. Res. **149**, 128 (1998).

[33] I. Tavernelli, U. F. Rohrig, and U. Rothlisberger, Mol. Phys. **103**, 963 (2005).

[34] I. Tavernelli *et al.*, Chem. Phys. Chem. **9**, 2099 (2008).

[35] J. de Vries, R. Hoekstra, R. Morgenstern, and T. Schlatholter, Phys. Rev. Lett. **91**, 053401 (2003).

[36] J. Tabet *et al.*, Phys. Rev. A **81**, 012711 (2010).

[37] B. Coupier *et al.*, Eur. Phys. J. D **20**, 459 (2002).

[38] J. de Vries, R. Hoekstra, R. Morgenstern, and T. Schlatholter, J. Phys. B: At. Mol. Opt. Phys **35**, 4373 (2002).

[39] J. de Vries, R. Hoekstra, R. Morgenstern, and T. Schlathoelter, Phys. Scr. **T110**, 336 (2004).

[40] L. Storchi *et al.*, J. Chem. Phys. **129**, 154309 (2008).

[41] O. Plekan *et al.*, Phys. Scr. **78**, 058105 (2008).

[42] M. Schwell, H.-W. Jochims, H. Baumgrtel, and S. Leach, Chem. Phys. **353**, 145 (2008).

[43] B. J. Alder and T. E. Wainwright, J. Chem. Phys. **27**, 1208 (1957).

[44] B. J. Alder and T. E. Wainwright, J. Chem. Phys. **31**, 459 (1959).

[45] A. Rahman, Phys. Rev. A **136**, 405 (1964).

[46] F. H. Stillinger and A. Rahman, J. Chem. Phys. **60**, 1545 (1974).

[47] J. A. McCammon, B. R. Gelin, and M. Karplus, Nature **267**, 585 (1977).

[48] J. M. Haile, *Molecular dynamics simulations: Elementary Methods* (Eilry Professional Paperback Series, New York, 1997).

[49] D. C. Rappaport, *The art of molecular dynamic simulation*, 2nd ed. (Cambridge University Press, UK, 2004).

[50] D. Frenkel and B. Smit, *Undestanding molecular simulation*, 2nd ed. (Academic press, California, 2002).

[51] A. R. Leach, *Molecular modeling: Principles and applications*, 2nd ed. (Pearson Education, England, 2001).

[52] G. Sutmann, in *Quantum simulation of complex many-body systems: from theory to algorithms, Lecture Notes*, Vol. 10 of *NIC Series*, edited by J. Grotendorst, D. Marx, and A. Maramatsu (Jon von Neumann Institute for Computing, ADDRESS, 2002), p. 211.

[53] P. Balbuena and J. M. Seminario, *From Classical to Quantum methods*, 1st ed. (Elsevier Science, Netherlands, 1999).

[54] M. E. Tuckerman, J. Phys. Chem. B **104**, 159 (2000).

[55] M. C. Payne *et al.*, Rev. Mod. Phys. **64**, 1045 (1992).

[56] R. Vuilleumier, Lect. Notes Phys. **703**, 217 (2006).

[57] *Multidimensional Quantum Dynamics*, 1st ed. (Wiley-VCH, Germany, 2009).

[58] *Introduction to Quantum Mechanics: a time-dependent perspective*, 1st ed. (University Science Books, California, 2007).

[59] E. Deumens, A. Diz, R. Longo, and Y. Ohrn, Rev. Mod. Phys. **66**, 917 (1994).

[60] *Introduction to Quantum Mechanics: a time-dependent perspective*, 1st ed. (Cambridge University Press, Cambridge, UK, 2009).

[61] J. C. Tully, in *Moder Methods for multidimensional dynamics computations in Chemistry*, edited by B. J. Berne, G. Ciccotti, and D. F. Coker (World Scientific, London, 1998), Chap. II.

[62] E. Schrodinger, Phys. Rev. **28**, 1049 (1926).

[63] D. S. Sholl and J. A. Steckel, *Density functional theory* (Cambridge University Press, New York, 2002).

[64] K. Gottfried, *Quantum Mechanics: Fundamentals* (Springer, New York, 2003).

[65] A. Messiah, *Quantum Mechanics*, 12th ed. (Dover Publications, INC., New York, 1991).

[66] P. Guttinger, Z. Phys. **73**, 169 (1932).

[67] W. Pauli, *Handbuch der Physik* (Springer, Berlin, 1933), p. 162.

[68] H. Hellmann, *Einfuhrung in die Quantenchemie* (Franz Deuticke, Leipzig, 1937), p. 285.

[69] R. Feynman, Phys. Rev. **56**, 340 (1939).

[70] P. Pulay, *Modern Theoretical Chemistry* (Plenum, New York, 1977), Vol. 4.

[71] B. M. Deb, Rev. Mod. Phys. **45**, 22 (1973).

[72] M. D. Ventra and S. T. Pantelides, Phys. Rev. B **61**, 16207 (2001).

[73] J. I. Steinfeld, *Molecules and Radiation: An Introduction to Modern Molecular Spectroscopy* (MIT Press, Cambridge, UK, 2005).

[74] M. Born and R. Oppenheimer, Annalen der Physik **84**, 457 (1927).

[75] J. Kohanoff, *Electronic structure calculations for solids and molecules*, 1st ed. (Cambridge, UK, 2006).

[76] M. Baer, *Beyond Born-Oppenheimer* (John Wiley and Sons, INC., New York, 2006), Chap. II.

[77] W. Domcke, D. R. Yarkony, and H.Koppel, *Canonical intersections* (World Scientific Publishing Co., London, 2004), Chap. I.

[78] P. A. M. Dirac, *Principles of Quantum Mechanics*, 3rd ed. (Oxford University Press, Oxford, 1947).

[79] D. Durr and S. Teufel, *Bohmian Mechanics: The Physics and Mathematics of Quantum Theory*, 1st ed. (Springer, Berlin, 2009).

[80] J. T. Bruer, Synthese **50**, 167 (1982).

[81] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3rd ed. (Addison Wesley, San Francisco, 2002).

[82] J. C. Tully, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, edited by B. J. Berne, G. Ciccotti, and D. F. Coker (World Scientific, Singapore, 1998), Chap. XXI.

[83] P. A. M. Dirac, Proc. Cambridge Philos. Soc. **26**, 376 (1930).

[84] E. J. Heller, J. Chem. Phys. **64**, 63 (1976).

[85] R. Harris, J. Chem. Phys. **72**, 1776 (1980).

[86] R. B. Gerber, V. Buch, and M. A. Ratner, J. Chem. Phys. **77**, 3022 (1982).

[87] N. Makri and W. H. Miller, J. Chem. Phys. **87**, 5781 (1987).

[88] R. B. Gerber and M. A. Ratner, Adv. Chem. Phys. **70**, 97 (1988).

[89] M. Messina and R. D. Coalson, J. Phys. Chem. **90**, 4015 (1989).

[90] K. Haug and H. Metiu, J. Phys. Chem. **99**, 6253 (1993).

[91] P. Ehrenfest, Zeitschrift fur Physik **45**, 455 (1927).

[92] N. L. Doltsinis and D. Marx, J. Theor. Comp. Chem. **2**, 319 (2002).

[93] J. Tully, M. Gomez, and M. Head-Gordon, J. Vac. Sci. Technol. **A11**, 1914 (1993).

[94] S. Klein *et al.*, Chem. Phys. Lett. **292**, 259 (1998).

[95] A. P. Horsfield, D. R. Bowler, and A. J. Fisher, J. Phys.: Condens. Matter **16**, L65 (2004).

[96] X. Li, J. C. Tully, H. B. Schlegel, and M. J. Frisch, J. Chem. Phys. **123**, 084106 (2005).

[97] C. M. Isborn, X. Li, and J. C. Tully, J. Chem. Phys. **126**, 134307 (2007).

[98] X. Andrade *et al.*, J. Chem. Theo. Comp. **5**, 728 (2009).

[99] W. Liang *et al.*, J. Phys. Chem. A **114**, 6201 (2010).

[100] R. Car and M. Parrinello, Phys. Rev. Lett. **22**, 2471 (1985).

[101] W. Koch, M. C. Holthausen, and E. J. Baerends, *A Chemist's Guide to Density Functional Theory* (WILEY-VCH, Weinheim, 2001).

[102] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Science **321**, 792 (2008).

[103] I. N. Levine, *Quantum chemistry* (Prentice Hall, New Jersey, 1999), Chap. I.

[104] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[105] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Berlin Heidelberg, Weinheim, 1990).

[106] W. Kohn and L. J. Sham, Phys. Rev **140**, A1133 (1965).

[107] R. G. Parr and W. Yang, *Density-functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).

[108] R. M. Martin, *Electronic Structure : Basic Theory and Practical Methods* (Cambridge University Press, ADDRESS, 2004).

[109] J. Slater, *Quantum Theory of Molecular and Solids. Vol. 4: The Self-Consistent Field for Molecular and Solids* (McGraw-Hill, New York, 1974).

[110] J. Slater, Phys. Rev. **81**, 385 (1951).

[111] S. Vowko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).

[112] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).

[113] U. von Barth, *Methods of Electronic Structure Calculations* (World Scientific, Singapure, 1994).

[114] F. Jensen, *Introduction to Computational Chemistry* (Wiley, New York, 1999).

[115] J. P. Perdew and W. Yang, Phys. Rev. B **33**, 8800 (1986).

[116] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).

[117] P. Gill, Mol. Phys. **89**, 433 (1996).

[118] J. Perdew, Phys. Rev. B **33**, 8822 (1986).

[119] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[120] R. Colle and O. Salvetti, Theor. Chim. Acta **37**, 4135 (1975).

[121] A. D. Corso, A. Pasquarello, A. Baldereschi, and R. Car, Phys. Rev. B **53**, 1180 (1996).

[122] M. Sprik, J. Hutter, and M. Parrinello, J. Chem. Phys. **105**, 1142 (1996).

[123] A. D. Becke, J. Chem. Phys. **104**, 1040 (1996).

[124] A. Alavi *et al.*, Phys. Rev. Lett. **80**, 3650 (1998).

[125] W. Pollard and R. Friesner, J. Chem. Phys. 99, 6742 (1993). **99**, 6742 (1993).

[126] A. Alavi, J. Kohanoff, M. Parrinello, and D. Frenkel, Phys. Rev. Lett. **73**, 2599 (1994).

[127] J. Cullum and R. A. Willoughby, *Lanczos algorithms for large symmetric eigenvalue computations, Vol. 1: Theory of Progress in Scientific Computing*, 3rd ed. (Birkhauser, Boston, 1985).

[128] M. P. Teter, M. C. Payne, and D. C. Allen, Phys. Rev. B **40**, 12255 (1989).

[129] P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).

[130] C. Csaszar and P. Pulay, J. Mol. Struc. **114**, 31 (1984).

[131] J. Hutter, H. P. Luthi, and M. Parrinello, Comput. Mat. Sci. **2**, 244 (1994).

[132] C.-O. Almbladh and U. von Barth, Phys. Rev. B **31**, 3231 (1985).

[133] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1st ed. (Brooks Cole, USA, 1976).

[134] M. P. M. Boero, K. Terakura, T. Ikeshoji, and C. C. Liew, Phys. Rev. Lett. **226403**, 90 (2003).

[135] C. Nicolas, R. Spezia, A. Boutin, and R. Vuilleumier, Phys. Rev. Lett. **91**, 208304 (2003).

[136] R. Spezia *et al.*, Mol. Sim. **30**, 749 (2004).

[137] J. C. Phillips, Phys. Rev. **112**, 685 (1958).

[138] M. L. Cohen and V. Heine, Solid State Phys. **24**, 37 (1970).

[139] M. T. Yin and M. L. Cohen, Phys. Rev. B **25**, 7403 (1982).

[140] P. Fuentealba, H. Preussa, H. Stolla, and L. V. Szentpály, Chem. Phys. Lett. **89**, 418 (1982).

[141] M. T. Yin and M. L. Cohen, Phys. Rev. B **52**, 11793 (1995).

[142] R. Hamann, M. Schluter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

[143] G. B. Bachelet, D. R. Hamann, and M. Schluter, Phys. Rev. Lett. **43**, 1494 (1979).

[144] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).

[145] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).

[146] D. A. Kleier, T. A. Halgren, J. H. H. Jr, and W. N. Lipscomb, J. Chem. Phys. **61**, 3905 (1974).

[147] J. M. Foster and S. F. Boys, Rev. Mod. Phys. **32**, 300 (1960).

[148] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).

[149] L. Verlet, Phys. Rev. **159**, 98 (1967).

[150] L. Verlet, Phys. Rev. **165**, 201 (1967).

[151] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics, Vol. 1*, 2nd ed. (Addison-Wesley, New-York, 2005).

[152] CPMD V3.13, copyright IBM Corp (1990-2008), copyright MPI fuer Festkoerperforschung Stuttgart (1997-2001).

[153] H. C. Andersen, J. Comp. Phys. **52**, 24 (1983).

[154] E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984).

[155] M. A. L. Marques and E. K. U. Gross, Annu. Rev. Phys. Chem. **55**, 427 (2004).

[156] M. A. L. Marques *et al.*, *Lecture Nothes in Physics: Time-Dependent Density Functional Theory* (Srpinger, Germany, 2005).

[157] K. Burke, J. Werschnik, and E. Gross, J. Chem. Phys. **123**, 062206 (2005).

[158] F. Furche and K. Burke, Annual Reports in Computational Chemistry **1**, 19 (2005).

[159] N. Maitra, A. Wasserman, and K. Burke, *What is time-dependent density functional theory? Successes and Challenges* (Springer, Germany, 2003).

[160] P. Elliott, K. Burke, and F. Furche, *Reviews in Computational Chemistry* (Wiley, Hoboken, NJ, 2009), p. 91.

[161] R. V. Leeuwen, Int. J. Mod. Phys. B **14**, 1969 (2001).

[162] M. E. Casida, *Recent Advances in Density Functional Methods* (World Scientific, Singapore, 1995), p. 155.

[163] R. van Leeuwen, Phys. Rev. Lett. **80**, 1280 (1998).

[164] L. V. Keldysh, Sov. Phys. **20**, 1018 (1965).

[165] J. Sakurai, *Modern Quantum Mechanics* (Addison Wesley Longman, USA, 1994), Chap. II.

[166] R. Baer and R. Gould, J. Chem. Phys. **114**, 3385 (2001).

[167] D. McQuarrie, *Statistical Mechanics* (Harper and Row, New York, 1976).

[168] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, 1987).

[169] R. E. Wilde and S. Singh, *Statistical Mechanics, Fundamentals and Modern Applications* (John Wiley and Sons, Inc, New York, 1998).

[170] L. V. Woodcock, Chem. Phys. Lett. **10**, 257 (1971).

[171] H. J. C. B. et al, J. Chem. Phys. **81**, 3684 (1984).

[172] S. Nosé, Mol. Phys. **53**, 255 (1984).

[173] W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).

[174] R. S. Mulliken, J. Chem. Phys. **23**, 1833 (1955).

[175] R. S. Mulliken, J. Chem Phys. **36**, 3428 (1962).

[176] A. E. Reed, L. A. Curtiss, and F. Weinhold, Chem. Rev. **88**, 899 (1988).

[177] P.-O. Lwdin, Adv. Quant. Chem. **5**, 185 (1970).

[178] R. F. W. Bader, *Atoms in molecules: A quantum theory* (Oxford University Press, Oxford, 1990).

[179] R. F. W. Bader, Chem. Rev **91**, 893 (1991).

[180] G. K. H. Madsen *et al.*, Phys. Rev. B **59**, 12359 (1999).

[181] R. F. W. Bader, Phys Rev B **49**, 13348 (1994).

[182] R. F. W. Bader, T. T. N. en Dang, and Y. Gal, Rep. Prog. Phys. **44**, 893 (1981).

[183] R. F. W. Bader, Y. Tal, S. G. Anderson, and T. T. N. en Dang, Israel J. Chem. **19**, 8 (1980).

[184] B. B. Stefanov and J. Cioslowski, J. Comput. Chem **16**, 1394 (1995).

[185] G. Henkelman, A. Arnaldsson, and H. Jnsson, Comp. Mat. Sci. **36**, 354 (2005).

[186] W. Tang, E. Sanville, and G. Henkelman, J. Phys.: Condens. Matter **20**, in press (2008).

[187] E. Sanville, S. D. Kenny, R. Smith, and G. Henkelman, J. Comput. Chem. **28**, 899 (2007).

[188] H.-J. Li and J.-J. Ho, J. Phys. Chem. C **114**, 1194 (2010).

[189] G. Giorgi, M. V. Schilfgaarde, A. Korkin, and K. Yamashita, Nanoscale. Res. Lett. **5**, 469 (2010).

[190] G. Giorgi, M. V. Schilfgaarde, A. Korkin, and K. Yamashita, J. Phys. Chem. A **114**, 4486 (2010).

[191] López-Tarifa *et al.*, submitted –, (2010).

[192] L. Adoui *et al.*, **48**, 075101 (2009).

[193] Gaussian 09 Revision A.1, gaussian Inc. Wallingford CT 2009.

[194] J. S. Binkley and J. A. Pople, J. Am. Chem. Soc. **102**, 939 (1980).

[195] W. J. Hehre, R. Ditchfiel, and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).

[196] R. Krishnan, J. S. Binkley, and R. S. an J. A. Pople, J. Chem. Phys. **72**, 650 (1980).

[197] M. J. Frisch, J. A. Pople, and J. S. Binkley, J. Chem. Phys. **80**, 3265 (1984).

[198] M. M. Francl *et al.*, J. Chem. Phys. **77**, 3654 (1982).

[199] B. Liu *et al.*, Phys. Rev. Lett. **97**, 133401 (2006).

[200] M. Nuevo *et al.*, Astrobiology **9**, 683 (2009).

[201] W. Tang, E. Sanville, and G. Henkelman, J. Phys.: Condens. Matter **20**, in press (2008).

[202] P. Cafarelli *et al.*, in *Radiation Damage in Biomolecular Systems*, edited by K. Tokési and B. Sulik (Springer, Debrecen, Hungary, 2008), Vol. 1080, p. 71.

[203] S. Denifl *et al.*, Int. J. Mass Spectrom. **238**, 47 (2004).

[204] H.-W. Jochims, M. Schwell, H. Baumgrtel, and S. Leach, Chem. Phys. **314**, 263 (2005).

[205] R. Vuilleumier and M. Sprik, J. Chem. Phys. **115**, 3454 (2001).

[206] M.-P. Gaigeot and M. Sprick, J. Phys. Chem. B **108**, 7458 (2004).

[207] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[208] S. Miertus, E. Scrocco, and J. Tomasi, J. Chem. Phys. **55**, 117 (1981).

[209] S. Miertus and J. Tomasi, J. Chem. Phys. **65**, 239 (1982).

[210] O. Tapia and J. Beltrán, *Solvent effects and chemical reactivity* (Kluwer Academic Publishers, New York, 2002).

[211] J. Tomasi and M. Persico, Chem. Rev. **94**, 2027 (1994).

[212] A. A. Varnek, G. Wipff, A. S. Glebov, and D. Feil, J. Comp. Chem. **16**, 1 (1995).

[213] M.-P. Gaigeot and M. Sprick, J. Phys. Chem. B **107**, 10345 (2003).

[214] J. R. Milligan *et al.*, Radiat. Res. **146**, 510 (1996).

[215] Z. Bacic and R. E. Miller, J. Phys. Chem. **100**, 12945 (1996).

[216] I. V. Hertel, C. Huglin, C. Nitsch, and C. P. Schulz, Phys. Rev. Lett **67**, 1767 (1991).

[217] F. N. Keutsch and R. J. Saykally, Proc. Natl. Acad. Sci. **98**, 10533 (2001).

[218] A. W. J. Castleman and K. H. J. Bowen, J. Phys. Chem. **100**, 12911 (1996).

[219] R. N. Barnett and U. Landman, Phys. Rev. Lett. **70**, 1775 (1993).

[220] O. Bjorneholm, F. Federmann, S. Kakar, and T. Moller, J. Chem. Phys. **111**, 546 (1999).

[221] Z.Sternovsky, M. Horányi, and S. Robertson, Phys. Rev. A **64**, 023203 (2001).

[222] F. H. Stillinger, Science **209**, 451 (1980).

[223] D. R. Miller, *Atomic and Molecular Beam Methods* (Oxford University Press, New York, 1988).

[224] J. Farges, M. F. de Feraudy, B. Raoult, and G. Torchet, J. Chem. Phys. **78**, 5067 (1983).

[225] C. Bobbert, S. Schutte, C. Steinbach, and U. Buck, Eur. Phys. J. D **19**, 183 (2002).

[226] S. Legendre, Master's thesis, University of Caen Basse-Normandie, 2006.

[227] S. Legendre *et al.*, J. Phys. B: At. Mol. Opt. Phys. **38**, L233 (2005).

[228] L. A. et al, Nucl. Instrum. Method B **245**, 94 (2006).

[229] L. A. et al, J. Phys: Conf. Ser. **88**, 012023 (2007).

[230] J.-P. G. C. Caraby, L. Adoui and A. Cassimi, Eur. Phys. J. D **2**, 53 (1998).

[231] M. T. et al, J. Phys. B: At. Mol. Opt. Phys. **33**, L11 (2000).

[232] G. Laurent *et al.*, Phys. Rev. Lett. **70**, 1775 (2006).

[233] O. Mó, M. Yaez, and J. Elguero, J. Chem. Phys. **97**, 1992 (6628).

[234] F. Dong, S. Heinbuch, J. J. Rocca, and E. R. Bernstein, J. Chem. Phys. **124**, 224319 (2006).

[235] F. Chandezon *et al.*, Phys. Rev. Lett. **87**, 153402 (2001).

[236] O. Echt, D. Kreisle, M. Knapp, and E. Recknagel, Chem. Phys. Lett. **108**, 401 (1984).

[237] G. M. Lancaster, F. Honda, Y. Fukuda, and J. W. Rabalais, J. Am. Chem. Soc. **101**, 1951 (1979).

[238] T. Matsuo *et al.*, J. Chem. Phys. **101**, 5356 (1994).

[239] J. M. H. et al, Science **308**, 1765 (2005).

[240] R. Ludwig, Chem. Phys. Chem. **5**, 1495 (2004).

[241] M. Eigen, Angew. Chem. **75**, 489 (1964).

[242] G. Zundel and H. Metzger, Z. Phys. Chem. **58**, 225 (1968).

[243] E. S. Wisnieswski, J. R. Stairs, and A. W. J. Castleman, Int. J. Mass. Spectrom. **212**, 273 (2001).

[244] T. Wroblewski, Radiat. Phys. Chem. **68**, 313 (2003).

[245] B. Manil *et al.*, Phys. Rev. Lett. **91**, 215504 (2003).

[246] V. Kumarappan, M. Krishnamurthy, and D. Mathur, Phys. Rev. A **67**, 063207 (2003).

[247] W. Dewey and R. Humphrey, Radiat. Res. **26**, 538 (1965).

[248] K. Sano, T. Hoshino, M. Nagai, and J. Neurosurg, Radiat. Res. **28**, 530 (1968).

[249] W. Pauli, Nobel Lecture (1946).

[250] A. Szabo and N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover Publications, INC., Mineola, New York, 1989).